

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using bw model

Run on: March 25, 2005, 22:00:09 ; Search time 9493.76 Seconds
(without alignments)
11376.592 Million cell updates/sec

Title: US-09-622-964A-2
Perfect score: 2229
Sequence: 1 caggagatccaccagccta.....aaaaaaaaaaaaaaaa 2229

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Genembl:*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_com:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_to:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2229	100.0	2229	6	BD136720 Best's ma
2	2208.4	99.1	2210	9	AF057169 Homo sapi
3	2166.2	97.2	2171	9	AF073501 Homo sapi
4	2155.8	96.7	2170	6	CQ729633 Sequence
5	2022.6	90.7	2187	9	AY357925 Macaca fa
6	2013	90.3	2429	6	BD136721 Best's ma
7	2002.4	89.8	2420	9	AF057170 Homo sapi
8	1793	80.4	2441	9	BC015320 Homo sapi
9	1756.4	78.4	1758	9	AY515704 Homo sapi
10	1754.8	78.7	1758	6	AX745964 Sequence
11	1584	71.1	2435	9	BC041664 Homo sapi
12	1247.2	56.0	1263	6	BD132659 Secreted
13	925.8	41.5	2035	10	BC079048 Ratcys no
14	893.4	40.1	1904	10	AY450427 Mus muscu
15	871.2	39.1	1916	6	BD136744 Best's ma
16	861.4	38.6	1285	6	AY064707 Sus scrof
17	642.2	28.8	16125	6	BD136719 Best's ma
18	642.2	28.8	142092	9	AF139813 Homo sapi
19	642.2	28.8	163024	9	AP066260 Homo sapi

c	20	642.2	28.8	163915	2	AC087451	AC087451 Homo sapi
	21	642.2	28.8	166867	9	AP003733	AP003733 Homo sapi
	22	642.2	28.8	196080	9	AC004228	AC004228 Homo sapi
	23	640.6	28.7	706	9	HSVMD2P10	AF073499 Homo sapi
	24	640.6	28.7	112309	9	AC003025	AC003025 Human chr
	25	640.6	28.7	133683	2	AC084857	AC084857 Homo sapi
	26	599	26.9	726	9	AY357926	AY357926 Macaca fa
	27	577.2	25.9	2137	6	AX745966	AX745966 Sequence
	28	576.8	25.9	1530	6	CQ731444	CQ731444 Sequence
	29	576.8	25.9	1530	6	AX745993	AX745993 Sequence
	30	576.8	25.9	1530	9	AY515705	AY515705 Homo sapi
	31	576.8	25.9	1908	9	AF440756	AF440756 Homo sapi
c	32	552.8	24.8	160169	2	AC051664	AC051664 Homo sapi
	33	538.4	24.2	1956	10	BC031186	BC031186 Mus muscu
	34	538.4	24.2	1957	10	BC019528	BC019528 Mus muscu
	35	538.4	24.2	1992	10	BC036157	BC036157 Mus muscu
	36	538.4	24.2	2028	10	BC036163	BC036163 Mus muscu
	37	535.2	24.0	1954	10	AY450428	AY450428 Mus muscu
	38	535	24.0	1861	9	AK000139	AK000139 Homo sapi
	39	494.4	22.2	1898	5	BC043854	BC043854 Xenopus l
	40	494.4	22.2	1921	5	AY273825	AY273825 Xenopus l
	41	484.6	21.7	1882	5	CR760914	CR760914 Xenopus t
	42	484.6	21.7	2021	5	BC061379	BC061379 Xenopus t
	43	478.4	21.5	2064	5	BC084229	BC084229 Xenopus l
	44	474	21.3	2719	10	AY450426	AY450426 Mus muscu
	45	472	21.2	1650	5	AY273826	AY273826 Xenopus l

ALIGNMENTS

RESULT 1	BD136720	2229 bp	DNA	linear	PAT 18-SEP-2002
LOCUS	BD136720				
DEFINITION	Best's macular dystrophy gene.				
ACCESSION	BD136720.1	GI:23231665			
VERSION	BD136720.1	GI:23231665			
KEYWORDS	JP 2002504559-A/2.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	Petrushkin,K., Caskey,T.C., Metzker,M. and Wadelius,C.				
TITLE	Best's macular dystrophy gene				
JOURNAL	Patent: JP 2002504559-A 2 12-FEB-2002;				
COMMENT	MERCK & CO INC, CLAES WADELIVS				
	OS Homo sapiens (human)				
	PN JP 2002504559-A/2				
	PP 22-FEB-1999 JP 2000533447				
	PR 25-FEB-1998 US 60/075941,18-DEC-1998 US 60/112926 PI				
	KONSTANTIN PETRUSHKIN,THOMAS C CASKEY,MICHAEL METZKER,CLAES PI				
	WADELIVS				
	PC C07K16/18,C07K14/47,C12N5/10,C12N15/09,C12P19/34,C12Q1/68// PC				
	C12P21/08,				
	PC C12N5/00,C12N15/00				
	CC Best's macular dystrophy gene				
	CH Key				
	FT source				
	FT 1..2229				
	FT /organism="Homo sapiens (human)".				
FEATURES	location/Qualifiers				
source	1..2229				
	/organism="Homo sapiens"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:9606"				
ORIGIN					
Query Match	100.0%;	Score 2229;	DB 6;	Length 2229;	
Best Local Similarity	100.0%;	Pred. No. 0;			
Matches 2229;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
oy	1	CAGGAGTCCACCAGCCTAGTCCAGACCTCTGTGGATCATCGACCCACCTGGAA	60		

```
Db 1 CAGGAGTCCCAAGCTAGTCCGAGACCTTCTGGGATCATGGACCACCTGGAA 60
Qy 61 CCCACCTGACCCCAAGCCCACTGCTGACGACCACTGCTGGCATGACATCATTA 120
Db 61 CCCACCTGACCCCAAGCCCACTGCTGACGACCACTGCTGGCATGACATCATTA 120
Qy 121 CAGCCAGAGTGGTAAATGCCCCGTTAGGCTCTTCCGCGCTGCTGTGTGGCGGG 180
Db 121 CAGCCAGAGTGGTAAATGCCCCGTTAGGCTCTTCCGCGCTGCTGTGTGGCGGG 180
Qy 181 GCGACATCTACAGAGCTGCTATATGCGAGATTCTTAACTTCTGCTCTGCTA 240
Db 181 GCGACATCTACAGAGCTGCTATATGCGAGATTCTTAACTTCTGCTCTGCTA 240
Qy 241 TCCGCTTTATTTATAGAGCTGCGCTCAGAGGAAACAAGAGTATGTTGAAACTGA 300
Db 241 TCCGCTTTATTTATAGAGCTGCGCTCAGAGGAAACAAGAGTATGTTGAAACTGA 300
Qy 301 CTCTGTATTTGCGACAGCTTACATCAAGCTCATCCCATTTCTTCTGCTGGCTTCA 360
Db 301 CTCTGTATTTGCGACAGCTTACATCAAGCTCATCCCATTTCTTCTGCTGGCTTCA 360
Qy 361 TGAAGCTGGTGTGACCCGCTGTGGAAACAGATAGAGAACTGCGTGGCCGACCG 420
Db 361 TGAAGCTGGTGTGACCCGCTGTGGAAACAGATAGAGAACTGCGTGGCCGACCG 420
Qy 421 TCATGAGCTGTGTCGGGCTTCGTGAAAGCAAGAGCAAGCAAGCGCGCTGGCG 480
Db 421 TCATGAGCTGTGTCGGGCTTCGTGAAAGCAAGAGCAAGCAAGCGCGCTGGCG 480
Qy 481 GCAAGCTCATCCGCTACCGCAACCTGGGCAAGCTCATCTGCGGACGCTACG 540
Db 481 GCAAGCTCATCCGCTACCGCAACCTGGGCAAGCTCATCTGCGGACGCTACG 540
Qy 541 CAGTATCAAGGCGCTTCCCAAGCGCCAGCACTGTGTGCAAGAGCTTTATGACTCG 600
Db 541 CAGTATCAAGGCGCTTCCCAAGCGCCAGCACTGTGTGCAAGAGCTTTATGACTCG 600
Qy 601 CAGAAACAAGAGTGGAGAACTGAGCTTACCAACAACATGTTCTGGGTGCGCTGG 660
Db 601 CAGAAACAAGAGTGGAGAACTGAGCTTACCAACAACATGTTCTGGGTGCGCTGG 660
Qy 661 TGTGTTTGGCCACTGTGCAATGAAAGGCTGTGGAGTGAATCCGGGACCTTATCC 720
Db 661 TGTGTTTGGCCACTGTGCAATGAAAGGCTGTGGAGTGAATCCGGGACCTTATCC 720
Qy 721 TGCTCAGAGCTGTGTAAGAGATGAACCTTGGCGTCACTGATGAGCACTGTATG 780
Db 721 TGCTCAGAGCTGTGTAAGAGATGAACCTTGGCGTCACTGATGAGCACTGTATG 780
Qy 781 CTTACGACTGGATTAGTATCCACTGTGTATATACAGAGTGTGACTGTGGCGGTGTA 840
Db 781 CTTACGACTGGATTAGTATCCACTGTGTATATACAGAGTGTGACTGTGGCGGTGTA 840
Qy 841 GCTTCTTCCGACCTTGTCTATGTTGGGCGGAGTTTCTGAACCCAGCAAGGCTTACC 900
Db 841 GCTTCTTCCGACCTTGTCTATGTTGGGCGGAGTTTCTGAACCCAGCAAGGCTTACC 900
Qy 901 GCGATGAGCTGAGCTGTGTTGCGCTTCAAGTTCTGAGTTCCTTCTTATGTTG 960
Db 901 GCGATGAGCTGAGCTGTGTTGCGCTTCAAGTTCTGAGTTCCTTCTTATGTTG 960
Qy 961 GCTGGCTGAAAGTGGCAGAGCACTCATCAACCCCTTGGAGAGATGATGATTTTG 1020
Db 961 GCTGGCTGAAAGTGGCAGAGCACTCATCAACCCCTTGGAGAGATGATGATTTTG 1020
Qy 1021 AGACCAACTGAGTTTGTGACAGGAATTTGCAAGGTTCCCTGTTGGCTGTGAGATGC 1080
Db 1021 AGACCAACTGAGTTTGTGACAGGAATTTGCAAGGTTCCCTGTTGGCTGTGAGATGC 1080
Qy 1081 ACCAGAGCTGCTCGATGAGCGGAGCATGTATGGAATTAAGCCGAGCCAGCCCC 1140
Db 1081 ACCAGAGCTGCTCGATGAGCGGAGCATGTATGGAATTAAGCCGAGCCAGCCCC 1140
Qy 1141 CTTACAGAGCTGTTCGCGCCAGATTCCGTGAGCTTCCTTTATGAGCTTCCACA 1200
Db 1141 CTTACAGAGCTGTTCGCGCCAGATTCCGTGAGCTTCCTTTATGAGCTTCCACA 1200
Qy 1201 TCAGCTGAAACAAAGAGAGATGAGTTTCAAGCCCAATCAGAGAGAGAGATGCTC 1260
Db 1201 TCAGCTGAAACAAAGAGAGATGAGTTTCAAGCCCAATCAGAGAGAGAGATGCTC 1260
Qy 1261 ACGCTGACATCATTTGCGCGCTTCTTGAAGCTGAGTCCCATGATCAATCTCCAG 1320
Db 1261 ACGCTGACATCATTTGCGCGCTTCTTGAAGCTGAGTCCCATGATCAATCTCCAG 1320
Qy 1321 CAAACTCAAGAGCCAACTATCTGTGGCCCAAGAGGAAATCCCTTCTCAAGAGGCTGC 1380
Db 1321 CAAACTCAAGAGCCAACTATCTGTGGCCCAAGAGGAAATCCCTTCTCAAGAGGCTGC 1380
Qy 1381 CCAAAAACCAAGGCGCCCAACAGAGGTTAGGGGCGAGAGACAAAGGCTTGA 1440
Db 1381 CCAAAAACCAAGGCGCCCAACAGAGGTTAGGGGCGAGAGACAAAGGCTTGA 1440
Qy 1441 AGCTTAAGCTGTGAGCGCTTCAAGTCTGGCCCATGATCAGAGCGAGCTACTA 1500
Db 1441 AGCTTAAGCTGTGAGCGCTTCAAGTCTGGCCCATGATCAGAGCGAGCTACTA 1500
Qy 1501 GTGCGCCAGAGAGCGCCCTTCAAGCTTCCATGTTCTTCCCTTGAACCATCAGCG 1560
Db 1501 GTGCGCCAGAGAGCGCCCTTCAAGCTTCCATGTTCTTCCCTTGAACCATCAGCG 1560
Qy 1561 CGTCAAGCTTCAAGATGTCAAGGATAGACCAAGAGCAAAAGCTTAAAGCTGTA 1620
Db 1561 CGTCAAGCTTCAAGATGTCAAGGATAGACCAAGAGCAAAAGCTTAAAGCTGTA 1620
Qy 1621 GTTCTGGGCGCAAGAAAGTTTGAATTCCTCTCAGAGAGCATGGGCTTGTATGAG 1680
Db 1621 GTTCTGGGCGCAAGAAAGTTTGAATTCCTCTCAGAGAGCATGGGCTTGTATGAG 1680
Qy 1681 ACCCAAGATATCTCAAGTGAAGAGGAAACGTGTGAGTTTAACTGAGGATATGCC 1740
Db 1681 ACCCAAGATATCTCAAGTGAAGAGGAAACGTGTGAGTTTAACTGAGGATATGCC 1740
Qy 1741 AGATCCCGCAAAATCACTTCAAGAAACCTTGTGAACAATCAACAACATACACTA 1800
Db 1741 AGATCCCGCAAAATCACTTCAAGAAACCTTGTGAACAATCAACAACATACACTA 1800
Qy 1801 CACTCAAAATCATATGATCTCTTATTTGGGCTTGGAAAACAGGATGAAGCAATTC 1860
Db 1801 CACTCAAAATCATATGATCTCTTATTTGGGCTTGGAAAACAGGATGAAGCAATTC 1860
Qy 1861 AACCTGCTTCTTAATGGGGATGCTTCCGCAAGCAAGTCTCACTGTGTGTAACACAGA 1920
Db 1861 AACCTGCTTCTTAATGGGGATGCTTCCGCAAGCAAGTCTCACTGTGTGTAACACAGA 1920
Qy 1921 GGAACATGATCCAGTCAAGCCATACAGCTGTGCACTGGAAGACGTGTCTTCAACAG 1980
Db 1921 GGAACATGATCCAGTCAAGCCATACAGCTGTGCACTGGAAGACGTGTCTTCAACAG 1980
Qy 1981 CCTGAATCAAAAGTTTACTTATATGATTAATAATCCAGACTTACGCTTTATG 2040
Db 1981 CCTGAATCAAAAGTTTACTTATATGATTAATAATCCAGACTTACGCTTTATG 2040
Qy 2041 TTTTATTTATTAATACTGTGAAGCTGAGCACTGAACATTTGAAACATTTATCAGACTC 2100
Db 2041 TTTTATTTATTAATACTGTGAAGCTGAGCACTGAACATTTGAAACATTTATCAGACTC 2100
Qy 2101 TGGATTCAAGTGGGAAACCTTATGTTTATCTGAATTCAGAGCAGCAGCTTAGAT 2160
Db 2101 TGGATTCAAGTGGGAAACCTTATGTTTATCTGAATTCAGAGCAGCAGCTTAGAT 2160
Qy 2161 ACTGCCCAACTATATGAGTTTAAATTAATAACAATCTCTTAAAAAATTTTTAAAA 2220
Db 2161 ACTGCCCAACTATATGAGTTTAAATTAATAACAATCTCTTAAAAAATTTTTAAAA 2220
```

QY 2221 AAAAAAAAA 2229
Db 2221 AAAAAAAAA 2229

RESULT 2
LOCUS AF057169
DEFINITION Homo sapiens beetrophin (VMD2) mRNA, alternatively spliced product.
ACCESSION AF057169
VERSION AF057169.1 GI:3335158
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 2210)
Petrukhin, K., Koist, M.J., Bakali, B., Li, W., Xie, G., Marknell, T., Berggren, O., Foreman, K., Holmgren, G., Andreassen, S., Vujic, M., Metzger, M.L., Caekey, C.T. and Wadelius, C.
Identification of the gene responsible for Beet macular dystrophy
Nat. Genet. 19 (3), 241-247 (1998)

TITLE
JOURNAL MEDLINE
PUBMED 9662395
REFERENCE 2 (bases 1 to 2210)
Petrukhin, K.
Direct Submission
Submitted (03-APR-1998) Human Genetics, Merck Research Laboratories, West Point, PA 19486, USA

FEATURES
SOURCE
1..2210
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q13"
1..2210
/gene="VMD2"
105..1862
/gene="VMD2"
/note="Beet's macular dystrophy gene; BMD; alternatively spliced isoform"
/codon_start=1
/product="beetrophin"
/protein_id="AAC64343.1"
/db_xref="GI:3335159"
translation="MTITVTSQVNAARLSPSRLLLQWRSIYKLAGEPLIFILCY IIRIVYALTEBOOLMEPEKLTICDSYIOQLPISPTGPTVTVTRMNOYENLPV PDRMLSLVSGFVEGDEQGRLLRLTLIYANGVLLIRSVTAHYKRFPSQHLVQA GFMTPEAKHOLEKLSLPHNMFVWVPMVWFSANLSKAWLIGRLDPLLQSLNENVTLLR TOCHLYAVDWISLPLVNTQVTVAVVSFPLTCLVGRFLNPAPAYPQHELDLVVPR FTLQFFVYGLKVAEQLINPEDDDDFETNMIVDRLLQVSLAVDEHMDLPRMEP DMWNKPEPOPPTAASAOFRASFMGSTFNI SINKEMEROPNOEDEDAHAGIGR FLGQSHDHHPRRANSRTKLMFKRESLHGLPKNHKAAIQNVRGQEDNKAHLKY DAFKAPLYQRPQYISAPQTPUSPTPMFPLPSAPSRLHSLVGTGIDTQDSLKIVSSG AKKSFELLSDEGALMEHPESQVRRKTVENLTDMPPIPNHLKEPLEOSPVIHTT LKDHMDPYWALENDEAHNS"

ORIGIN
Query Match 99.1%; Score 2208.4; DB 9; Length 2210;
Beet Local Similarity 100.0%; Pred. No. 0;
Matches 2209; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAGGAGTCCCAACGAGCTTGTGCGAGACCTTGTGGATCATCGAGCCACCTGGAA 60
Db 1 CAGGAGTCCCAACGAGCTTGTGCGAGACCTTGTGGATCATCGAGCCACCTGGAA 60

QY 61 CCCACCTGACCAAGCCACCTGCTGAGCCGACCTGCGCTGCGCATGACCATCTTACA 120
Db 61 CCCACCTGACCAAGCCACCTGCTGAGCCGACCTGCGCTGCGCATGACCATCTTACA 120

QY 121 CAAGCAAGTGGCTAATGCCCGCTTAAAGCTCTTCTCCCGCTGCTGCTGCGGG 180
Db 121 CAAGCAAGTGGCTAATGCCCGCTTAAAGCTCTTCTCCCGCTGCTGCTGCGGG 180

QY 181 GCAGCATCTTACAGCTGCTATATGCGAGATCTTAAATCTTCTGCTGCTACTACATCA 240
Db 181 GCAGCATCTTACAGCTGCTATATGCGAGATCTTAAATCTTCTGCTGCTACTACATCA 240

QY 241 TCCGCTTATTTAATAGCTGCGCTCAACGGAAGAACAAACACTGATGTGGAAGACGA 300
Db 241 TCCGCTTATTTAATAGCTGCGCTCAACGGAAGAACAAACACTGATGTGGAAGACGA 300

QY 301 CTCTGATTTGCGAGAGCTATCCAGCTCATCCCATTTCTTGTGCTGCGGCTTCAAG 360
Db 301 CTCTGATTTGCGAGAGCTATCCAGCTCATCCCATTTCTTGTGCTGCGGCTTCAAG 360

QY 361 TGAAGCTGTGCTGAGACCGCTGCTGGAACAGTACGAGAACCTGCGGCGGACCGC 420
Db 361 TGAAGCTGTGCTGAGACCGCTGCTGGAACAGTACGAGAACCTGCGGCGGACCGC 420

QY 421 TCATGAGCTGTGCTGCGGCTTCTGCAAGGCAAGACGAGCAAGCCGCGCTGCGGC 480
Db 421 TCATGAGCTGTGCTGCGGCTTCTGCAAGGCAAGACGAGCAAGCCGCGCTGCGGC 480

QY 481 GCAGCTCATCCGCTAGCGCCAACTGAGGCAAGTGTCTATCTGCGAGCTCAGACCG 540
Db 481 GCAGCTCATCCGCTAGCGCCAACTGAGGCAAGTGTCTATCTGCGAGCTCAGACCG 540

QY 541 CAGTCTACAAAGCGTTTCCCGAGCGCCAGCACTGTGCAAGAGCTTTATGACTCCGG 600
Db 541 CAGTCTACAAAGCGTTTCCCGAGCGCCAGCACTGTGCAAGAGCTTTATGACTCCGG 600

QY 601 CAGAACCAAGCACTTGGAGAACTGAGCCCTACCAACAAATTTCTGGGTGCGCGGG 660
Db 601 CAGAACCAAGCACTTGGAGAACTGAGCCCTACCAACAAATTTCTGGGTGCGCGGG 660

QY 661 TGTGTTTGGCCAACTGTCATGAAAGCGTGGTGGAGTCCGAGACCTTATCC 720
Db 661 TGTGTTTGGCCAACTGTCATGAAAGCGTGGTGGAGTCCGAGACCTTATCC 720

QY 721 TGTCTCAGAGCTGCTGAGAGAGATGAACACTTGTGCTACTCAGTGTGACACTGTATG 780
Db 721 TGTCTCAGAGCTGCTGAGAGAGATGAACACTTGTGCTACTCAGTGTGACACTGTATG 780

QY 781 CCTACGACTGATTAATATCCACTGCTGTATTAACAGAGTGTACTGTGCGGTATCA 840
Db 781 CCTACGACTGATTAATATCCACTGCTGTATTAACAGAGTGTACTGTGCGGTATCA 840

QY 841 GCTTCTTCTGACTTGTCTAGTGTGGGCGGAGTTTGAACCCAGCCAGGCTTACCTG 900
Db 841 GCTTCTTCTGACTTGTCTAGTGTGGGCGGAGTTTGAACCCAGGCTTACCTG 900

QY 901 GCATGAGCTGAGACTGCTGTGTGCGCGTCTTCAAGTTCCTGCAATGATGTTG 960
Db 901 GCATGAGCTGAGACTGCTGTGTGCGCGTCTTCAAGTTCCTGCAATGATGTTG 960

QY 961 GCTGCTGAGAGTGGCAGAGCACTCAACCCCTTTGGAGAGATGATATATATTTTG 1020
Db 961 GCTGCTGAGAGTGGCAGAGCACTCAACCCCTTTGGAGAGATGATATATATTTTG 1020

QY 1021 AGACCAACTGGAATGTGAGAGAGAAATTTGAGAGTGTGCTGCTGATGAGATGC 1080
Db 1021 AGACCAACTGGAATGTGAGAGAGAAATTTGAGAGTGTGCTGCTGATGAGATGC 1080

QY 1081 ACCAGAGACTGCTCTGAGATGAGCGGACATGTACTGGAATTAAGCCGAGCCAGCC 1140
Db 1081 ACCAGAGACTGCTCTGAGATGAGCGGACATGTACTGGAATTAAGCCGAGCCAGCC 1140

QY 1141 CCTTCAAGCTGCTTCCGCGCAGTTCGCTGAGAGCTCTTATATGGGCTCCACTTCAACA 1200
Db 1141 CCTTCAAGCTGCTTCCGCGCAGTTCGCTGAGAGCTCTTATATGGGCTCCACTTCAACA 1200

QY 1201 TCAGCCTGAACAAAGAGAGATGATGATCCAGCCCAATCATGAGAGAGAGAGATGCTC 1260
Db 1201 TCAGCCTGAACAAAGAGAGATGATGATCCAGCCCAATCATGAGAGAGAGAGATGCTC 1260
QY 1261 ACCGTGGCATCATTTGGCCGCTTCCCTAGGCTGCAATGATCCATGATCCATCCAGAG 1320
Db 1261 ACCGTGGCATCATTTGGCCGCTTCCCTAGGCTGCAATGATCCATGATCCATCCAGAG 1320
QY 1321 CAAACTCAAGAGACCAAACTACTGTGGCCCAAGAGAGAAATCCCTCTCCCAAGAGAGAG 1380
Db 1321 CAAACTCAAGAGACCAAACTACTGTGGCCCAAGAGAGAAATCCCTCTCCCAAGAGAGAG 1380
QY 1381 CCAAAAACCAAG 1440
Db 1381 CCAAAAACCAAG 1440
QY 1441 AGCTTAAGGCTGAG 1500
Db 1441 AGCTTAAGGCTGAG 1500
QY 1501 GTGGCCCAAG 1560
Db 1501 GTGGCCCAAG 1560
QY 1561 CGTCAAAAGCTTCAAG 1620
Db 1561 CGTCAAAAGCTTCAAG 1620
QY 1621 GTTCTGGGGGCAAG 1680
Db 1621 GTTCTGGGGGCAAG 1680
QY 1681 ACCCAAGAGATCTCAAG 1740
Db 1681 ACCCAAGAGATCTCAAG 1740
QY 1741 AGATCCCGGAAATCACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
Db 1741 AGATCCCGGAAATCACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
QY 1801 CACTCAAAAGATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
Db 1801 CACTCAAAAGATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
QY 1861 AACCGGCTTCTCAAG 1920
Db 1861 AACCGGCTTCTCAAG 1920
QY 1921 GAGACCTGATCAAG 1980
Db 1921 GAGACCTGATCAAG 1980
QY 1981 CCTGAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2040
Db 1981 CCTGAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2040
QY 2041 TTTTATTCATTAATAAAGCTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2100
Db 2041 TTTTATTCATTAATAAAGCTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2100
QY 2101 TGGATTCAAG 2160
Db 2101 TGGATTCAAG 2160
QY 2161 ACTGCCCAAACTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2210
Db 2161 ACTGCCCAAACTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2210

RESULT 3
AF073501 2171 bp mRNA linear PRI 02-SEP-1998
LOCUS Homo sapiens vitelliform macular dystrophy protein (VMD2) mRNA,
DEFINITION

complete cds.
AF073501
AF073501.1 GI:3511241
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euteleostomi; Primates; Catarrhini; Homnidae; Homo.
Stohr,H., Marguard,A., Rivera,A., Cooper,P.R., Nowak,N.J.,
Shows,T.B., Gerhard,D.S. and Weber,B.H.
A gene map of the Best's vitelliform macular dystrophy region in
chromosome 11q12-q13.1
Genome Res. 8 (1), 48-56 (1998)
98112782
MEDLINE
PUBMED
2 (bases 1 to 2171)
Marguard,A., Stohr,H., Passmore,L., Kraemer,F., Rivera,A. and
Weber,B.H.F.
Direct Submission
Submitted (22-JUN-1998) Human Genetics, University, Biozentrum, Am
Hudland, Wuerzburg 97074, Germany
Location/Qualifiers
1..2171
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q12-q13.1"
/tissue_type="retinal pigment epithelium"
1..2171
/gene="VMD2"
89..1846
/gene="VMD2"
/note="Best disease"
/codon_start=1
/product="vitelliform macular dystrophy protein"
/protein_id="AAC33766.1"
/db_xref="GI:3511242"
/translation="MTITYTSGVNAARLGSRSRLLCMRGSIYKLLYSEFLIFELCY
IRFIRYRLATREQQLMERKTLVCDSTIQLPISFVIGFVTVLVTVMMQYENLP
PRLMSLVSGFVEKDEQRLRLRLRYANLGNVLIRSVTAHYKSPSAOHLYVA
GMTPAEHKOELKSLPHNMFVWVWPMFANLSMKAMLGRIIDPILLOSLINEMTLR
TQGHLYAVDWISIPLYTVQVTVAVSFFFLICLVRQFLNPAPAKYPGHELDLVVVF
TELQFFVYGLKVAEOLINPGEEDDDFEETNIVDRNLQVSLAVDEMDHDLPRME
DWMKRPBPQPYTAASQFRAFSMSSTPMISLKNEMEMQPNQDEEDHAAGTIGR
PLGLSHDHHPRAHSRKLMPKRESLILHGLPKNHAKQNVRGQDNKAMLUKAV
DAFKSAPLYKRPQYISAPQTPUSPTWFFPLEPSAPSKLHVTGIDTKDKSLKTVSSG
AKKSFELSESDDGALMEHPVSVQVRKRTVEFNLTDMPEIPNHKLEPLEOSPVIHTT
IKDHMDPYWALENDRDEAHS"

ORIGIN

Query Match 97.2%; Score 2166.2; DB 9; Length 2171;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2168; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 17 CCTAGTGCAGACCTTCTGTGATCATCGAAGCCCACTGAAGCCCACTGAAGCCCAAG 76
Db 1 CCTAGTGCAGACCTTCTGTGATCATCGAAGCCCACTGAAGCCCACTGAAGCCCAAG 60
QY 77 CCCACCTGCTGAGCCCACTGCTGCGCATGACATCATTAACAAGCCCAAGTGGCTAA 136
Db 61 CCCACCTGCTGAGCCCACTGCTGCGCATGACATCATTAACAAGCCCAAGTGGCTAA 120
QY 137 TGCCCGCTTAGAGCTCTTCTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 196
Db 121 TGCCCGCTTAGAGCTCTTCTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
QY 197 GCTATATGCGAGATTCTTAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 256
Db 181 GCTATATGCGAGATTCTTAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
QY 257 GCTGCGCTTACGGAAGAACAGCTGATGTTTGAAGAACTGACTGTGATTTGGCCAG 316

Db 241 GCTGGCCCTCA CGAAGAACAA CAGCTGATGTTTGAGAAATGACTGTGATTCGACAG 300
Qy 317 CTATCATCAGCTCATCCCATATTTCTCTGCTGCTGGCTTTCTACTGACGCTGCTGAC 376
Db 301 CTATCATCAGCTCATCCCATATTTCTCTGCTGCTGGCTTTCTACTGACGCTGCTGAC 360
Qy 377 CCGCTGTGGAACAGTACGAAAGCTGCGTGGCCCGACCGCTCTCAATGAGCTGTGTC 436
Db 361 CCGCTGTGGAACAGTACGAAAGCTGCGTGGCCCGACCGCTCTCAATGAGCTGTGTC 420
Qy 437 GGGCTTGTGGAAGCAAGACGAGCAAGCGCGCTGCGGCGCAGCTTCATCCGCTA 496
Db 421 GGGCTTGTGGAAGCAAGACGAGCAAGCGCGCTGCGGCGCAGCTTCATCCGCTA 480
Qy 497 CGCCAACTGGGCAACGTGCTCATCTGCGAGAGGTCAAGCAACCGCATGTCAGAGCCTT 556
Db 481 CGCCAACTGGGCAACGTGCTCATCTGCGAGAGGTCAAGCAACCGCATGTCAGAGCCTT 540
Qy 557 CCCCAGGCGCCAGCAGCTGTGCAAGCAGCTTATGACTCCGCGAAGCAAGCAGTT 616
Db 541 CCCCAGGCGCCAGCAGCTGTGCAAGCAGCTTATGACTCCGCGAAGCAAGCAGTT 600
Qy 617 GAGAAACTGAGCCTACACACAAATGTTCTGGGTGCGCTGGGTGCTTTCGCAACT 676
Db 601 GAGAAACTGAGCCTACACACAAATGTTCTGGGTGCGCTGGGTGCTTTCGCAACT 660
Qy 677 GTCAATGAGAGCGTGGCTGTGAGGTGCAATCGGGAACCTTATCTGCTTCAGAGCCTGCT 736
Db 661 GTCAATGAGAGCGTGGCTGTGAGGTGCAATCGGGAACCTTATCTGCTTCAGAGCCTGCT 720
Qy 737 GAAAGATGAAACACTTGGGCTGACTCACTGATGGAACACTGATGACTCAAGTGGATTAG 796
Db 721 GAAAGATGAAACACTTGGGCTGACTCACTGATGGAACACTGATGACTCAAGTGGATTAG 780
Qy 797 TATCCACTGCTGTATACACAGTGTGATCTGTGGCGGTGTACAGCTTCTTCTGACTTG 856
Db 781 TATCCACTGCTGTATACACAGTGTGATCTGTGGCGGTGTACAGCTTCTTCTGACTTG 840
Qy 857 TCTAGTTGGGGGCGAGTTTCTGAAACCGACCGAAGCTTACCTGCGCATGACTGACCT 916
Db 841 TCTAGTTGGGGGCGAGTTTCTGAAACCGACCGAAGCTTACCTGCGCATGACTGACCT 900
Qy 917 CGTTGTGCGGCTTCAAGTCTTCAAGTCTTCTGATGTTGGCTGGAGAGTGGC 976
Db 901 CGTTGTGCGGCTTCAAGTCTTCAAGTCTTCTGATGTTGGCTGGAGAGTGGC 960
Qy 977 AGACAGCTCATCAACCCCTTGGAGAGATGATGATTTTGACACCACTGGATTGT 1036
Db 961 AGACAGCTCATCAACCCCTTGGAGAGATGATGATTTTGACACCACTGGATTGT 1020
Qy 1037 CGACAGAAATTTGCAAGTGTCTCTGTTGGCTGTGAGTGAATGACACGACCTGCTCG 1096
Db 1021 CGACAGAAATTTGCAAGTGTCTCTGTTGGCTGTGAGTGAATGACACGACCTGCTCG 1080
Qy 1097 GATGGAACCGGACATGATCTGGAATPAAGCCGAGCCACAGCCCTTCAAGTGTCTC 1156
Db 1081 GATGGAACCGGACATGATCTGGAATPAAGCCGAGCCACAGCCCTTCAAGTGTCTC 1140
Qy 1157 CGCCCAATTCGCTGAGAGCTCTTATGAGGCTTCAACCTTCAACATCAACCTGAACAAAGA 1216
Db 1141 CGCCCAATTCGCTGAGAGCTCTTATGAGGCTTCAACCTTCAACATCAACCTGAACAAAGA 1200
Qy 1217 GGAATGGAATTTCAAGCCCAATCAAGAGAGCAGAGAGAGTCTCAAGCTGGCATATTGG 1276
Db 1201 GGAATGGAATTTCAAGCCCAATCAAGAGAGCAGAGAGAGTCTCAAGCTGGCATATTGG 1260
Qy 1277 CCGCTTCTAGGCGCTGAGAGTCCCATGATCAACATCTTCCAGGGCAAACTGAAGACCA 1336
Db 1261 CCGCTTCTAGGCGCTGAGAGTCCCATGATCAACATCTTCCAGGGCAAACTGAAGACCA 1320
Qy 1337 ACTACTGTGCGCAAGAGGGAATCCCTTCTCAGAGGCGCTGCGCAAAACCAAGAGC 1396

Db 1321 ACTACTGTGCGCAAGAGGGAATCCCTTCTCAGAGGCGCTGCGCAAAACCAAGAGC 1380
Qy 1397 AGCCAAACAGAACGTTAGGGGCGCAGAAAGCAACCAAGGCTTGAAGGCTGTGGA 1456
Db 1381 AGCCAAACAGAACGTTAGGGGCGCAGAAAGCAACCAAGGCTTGAAGGCTGTGGA 1440
Qy 1457 CGCCTTCAAGTCTGAGCCCACTGTATCAAGAGCCAGGCTTACTACAGTCCCCCAAGAGCC 1516
Db 1441 CGCCTTCAAGTCTGAGCCCACTGTATCAAGAGCCAGGCTTACTACAGTCCCCCAAGAGCC 1500
Qy 1517 CTTGAGCCCACTTCCCATGTTCTTCCCTTGAACCATCAAGCCGCTTCAAGCTTCAAG 1576
Db 1501 CTTGAGCCCACTTCCCATGTTCTTCCCTTGAACCATCAAGCCGCTTCAAGCTTCAAG 1560
Qy 1577 TGTACAGGATAGACACCAAGCAAAAGCTTAAAGACTGTGAGTTCTGAGGCGCAAGAA 1636
Db 1561 TGTACAGGATAGACACCAAGCAAAAGCTTAAAGACTGTGAGTTCTGAGGCGCAAGAA 1620
Qy 1637 AAGTTTGAATTTGCTCTCAGAGAGCGATGGGCGCTTGTATGAGAGCACCAGAGATCTCA 1696
Db 1621 AAGTTTGAATTTGCTCTCAGAGAGCGATGGGCGCTTGTATGAGAGCACCAGAGATCTCA 1680
Qy 1697 AGTGAAGAGAAACTGTGAGATTAACTGACGAGATATGCGAGATATCCCGAAATCA 1756
Db 1681 AGTGAAGAGAAACTGTGAGATTAACTGACGAGATATGCGAGATATCCCGAAATCA 1740
Qy 1757 CTTCAAGAACTTTTGGAAATTCACCAACATCACTACACTACATCAAGATCAAT 1816
Db 1741 CTTCAAGAACTTTTGGAAATTCACCAACATCACTACACTACATCAAGATCAAT 1800
Qy 1817 GSAATCTTATTTGGGCTTGGAAAAACAGGATGAAAGCATTTCTTCACTGCTTCTATG 1876
Db 1801 GSAATCTTATTTGGGCTTGGAAAAACAGGATGAAAGCATTTCTTCACTGCTTCTATG 1860
Qy 1877 GGGATGCTTGGCCAGCCAGAGTCTCACTGTGTGTACACAGCAGAACATGATCCAGTC 1936
Db 1861 GGGATGCTTGGCCAGCCAGAGTCTCACTGTGTGTACACAGCAGAACATGATCCAGTC 1920
Qy 1937 ACAGCCATAGAGCTGTACACATGAAAGAGTGTCTTCAACACAGCCGAATCAATGGTT 1996
Db 1921 ACAGCCATAGAGCTGTACACATGAAAGAGTGTCTTCAACACAGCCGAATCAATGGTT 1980
Qy 1997 AGCTTAATAGTAAATAATCCAGACTACTTCAAGCTTAAATGCTTATTCATTAATAAC 2056
Db 1981 AGCTTAATAGTAAATAATCCAGACTACTTCAAGCTTAAATGCTTATTCATTAATAAC 2040
Qy 2057 TGTGAAGCTTGAAGCTGAACCAATTTGAACATTTAACTCAAGCTGTGATTCAGAGTGGG 2116
Db 2041 TGTGAAGCTTGAAGCTGAACCAATTTGAACATTTAACTCAAGCTGTGATTCAGAGTGGG 2100
Qy 2117 AACCTTAAGTTTATCTGAATCCAAAGACGCCACCTTGTATGATGAGCCCAACTAATG 2176
Db 2101 AACCTTAAGTTTATCTGAATCCAAAGACGCCACCTTGTATGATGAGCCCAACTAATG 2160
Qy 2177 AGTTTAATAA 2187
Db 2161 AGTTTAATAA 2171

RESULT 4
CQ729633 2170 bp DNA linear PAT 03-FEB-2004
LOCUS CQ729633
DEFINITION Sequence 15567 from Patent WO02068579.
ACCESSION CQ729633
VERSION CQ729633.1 GI:42301252
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Bukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
Kits, such as nucleic acid arrays, comprising a majority of

humanexons or transcripts, for detecting expression and other uses thereof
Patent: WO 02068579-A, 15567 06-SEP-2002;
PE Corporation (NY) (US)

FEATURES
Source Location/Qualifiers
1..2170

ORIGIN
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

Query Match 96.7%; Score 2155.8; DB 6; Length 2170;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2168; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 17 CCTAGTGGCCAGACCTTCTGTGGATGATGGAACCCACCTGGAACCCACCTGACCCAAAG 76
Db 1 CCTAGTGGCCAGACCTTCTGTGGATGATGGAACCCACCTGGAACCCACCTGACCCAAAG 60
QY 77 CCCACCTGTGCAGACCCACCTGTGGCCATGACCATCACTTACAGAACGCAAGTGGCTTA 136
Db 61 CCCACCTGTGCAGACCCACCTGTGGCCATGACCATCACTTACAGAACGCAAG-GGCTTA 119
QY 137 TGCCTGGCTTAAAGCTCTTCCCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 196
Db 120 TGCCTGGCTTAAAGCTCTTCCCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 179
QY 197 GCTAATATGGGAGTCTTAAATCTTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 256
Db 180 GCTAATATGGGAGTCTTAAATCTTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 239
QY 257 GCTGGGCCCTCAAGGAGAAACAAGCTGATGTTTGAAGAACTGACTGTGTATTTGGACAG 316
Db 240 GCTGGGCCCTCAAGGAGAAACAAGCTGATGTTTGAAGAACTGACTGTGTATTTGGACAG 299
QY 317 CTACATCCAGCTTCATCCCATTTCTTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 376
Db 300 CTACATCCAGCTTCATCCCATTTCTTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 359
QY 377 CGCGTGTGGAGAACAGTATGAGAAACCTGCGGTGCGGACCGGCTCATGAGCTGTGTGTC 436
Db 360 CGCGTGTGGAGAACAGTATGAGAAACCTGCGGTGCGGACCGGCTCATGAGCTGTGTGTC 419
QY 437 GGGCTTTCGTGAGAGCAAGAGCAGACGAGCGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 496
Db 420 GGGCTTTCGTGAGAGCAAGAGCAGACGAGCGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 479
QY 497 CGCGAATCTGTGGCAAGTGTCTATCTGTGCGAGCGTCAAGCAGCTTCAAGCGCTT 556
Db 480 CGCGAATCTGTGGCAAGTGTCTATCTGTGCGAGCGTCAAGCAGCTTCAAGCGCTT 539
QY 557 CCCGAGCGCCAGCAGCTGTGTGCAAGCAGCTTATATCTCCGAGAGAACAGAGATT 616
Db 540 CCCGAGCGCCAGCAGCTGTGTGCAAGCAGCTTATATCTCCGAGAGAACAGAGATT 599
QY 617 GGAAGAACTGAGCTTACCAACAACATGTTCTGTGGTCCCTGTGGTGTGTTTSCAAACCT 676
Db 600 GGAAGAACTGAGCTTACCAACAACATGTTCTGTGGTCCCTGTGGTGTGTTTSCAAACCT 659
QY 677 GTCAATGAAGCGGTGTGAGAGTGAATCCGGAACCTTATCTGTCTCAAGAGCTGTCT 736
Db 660 GTCAATGAAGCGGTGTGAGAGTGAATCCGGAACCTTATCTGTCTCAAGAGCTGTCT 719
QY 737 GAACGAGATGAACAACCTTGTGCTACTAGTGTGAGACCTGTATGCTTACAGCTGAGTTAG 796
Db 720 GAACGAGATGAACAACCTTGTGCTACTAGTGTGAGACCTGTATGCTTACAGCTGAGTTAG 779
QY 797 TATCCACATGAGTGTATACAGAGTGTGACTGTGCGGTGTACAGCTTCTTCTGACTTGG 856
Db 780 TATCCACATGAGTGTATACAGAGTGTGACTGTGCGGTGTACAGCTTCTTCTGACTTGG 839
QY 857 TCTAGTTGGCGGAGCTTCTGAAACCAAGCCAAAGGCTTACCTGTGCGATGAGTGGAGACT 916
|||||

Db 840 TCTAGTTGGCGGAGCTTCTGAAACCAAGCCCTACCTTGGCCATGAGCTGACCT 899
QY 917 CGTTTGGCCCGCTCTTCAAGTTCCTGCAAGTCTTCTTATGTTGGCTGTGAGAGTGGC 976
Db 900 CGTTTGGCCCGCTCTTCAAGTTCCTGCAAGTCTTCTTATGTTGGCTGTGAGAGTGGC 959
QY 977 AGAGCAGCTCATCAACCCCTTGTGAGAGGATGATGATGATTTTGAACCAACTGAGATTGT 1036
Db 960 AGAGCAGCTCATCAACCCCTTGTGAGAGGATGATGATGATTTTGAACCAACTGAGATTGT 1019
QY 1037 GCAAGGAATTTGCAAGTGTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1096
Db 1020 GCAAGGAATTTGCAAGTGTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1079
QY 1097 GATGAGACCGGACATGATACGAAATTAACCCGAGCCCAAGCCCCCTTACAGCTGTCTTC 1156
Db 1080 GATGAGACCGGACATGATACGAAATTAACCCGAGCCCAAGCCCCCTTACAGCTGTCTTC 1139
QY 1157 CGCCAGTTTCGTGAGACCTCTTTATGAGCTTCACTTCAACATGACGCTGAAACAAAGA 1216
Db 1140 CGCCAGTTTCGTGAGACCTCTTTATGAGCTTCACTTCAACATGACGCTGAAACAAAGA 1199
QY 1217 GGAATGGAATTCAGACCCATATGAGAGACGAGAGATGCTCAAGCTGTGATCTTGG 1276
Db 1200 GGAATGGAATTCAGACCCATATGAGAGACGAGAGATGCTCAAGCTGTGATCTTGG 1259
QY 1277 CGGCTTCTAAGCTGACAGTCCATGATCAGCATCTCCAGAGGAACTCAAGAGCAAA 1336
Db 1260 CGGCTTCTAAGCTGACAGTCCATGATCAGCATCTCCAGAGGAACTCAAGAGCAAA 1319
QY 1337 ACTACTGTGGCCCAAGAGGAAATCCCTTCTCCAGAGGCTGTCCCAAAACCAAGAGC 1396
Db 1320 ACTACTGTGGCCCAAGAGGAAATCCCTTCTCCAGAGGCTGTCCCAAAACCAAGAGC 1379
QY 1397 AGCCAAACAGAACTTTAAGGGCCAGAGAGACAAACAGGCTGTGAAGCTTAAAGCTGTGGA 1456
Db 1380 AGCCAAACAGAACTTTAAGGGCCAGAGAGACAAACAGGCTGTGAAGCTTAAAGCTGTGGA 1439
QY 1457 CGCTTCAAGTGTGGCCCACTGTATCAGAGGCTCAGGCTTCAAGTGTGCTGTGCTGTGCTGTGCT 1516
Db 1440 CGCTTCAAGTGTGGCCCACTGTATCAGAGGCTCAGGCTTCAAGTGTGCTGTGCTGTGCTGTGCT 1499
QY 1517 CCTCAGCCCACTCCCAATGTTCTTCCCTGTGAACCAATGAGCGGCTTCAAGCTTCCAG 1576
Db 1500 CCTCAGCCCACTCCCAATGTTCTTCCCTGTGAACCAATGAGCGGCTTCAAGCTTCCAG 1559
QY 1577 TGTCAAGGCTATGACACCAAGCAAAAGCTTAAAGACTGTGAGTTCTGAGGCGCAAGAA 1636
Db 1560 TGTCAAGGCTATGACACCAAGCAAAAGCTTAAAGACTGTGAGTTCTGAGGCGCAAGAA 1619
QY 1637 AAGTTTGAATTTGCTCAGAGAGCGATGGGCTTGAATGAGACATCCAGAGATATCTCA 1696
Db 1620 AAGTTTGAATTTGCTCAGAGAGCGATGGGCTTGAATGAGACATCCAGAGATATCTCA 1679
QY 1697 AGTGAAGAGAAACTGTGAGATTAACTTGAAGATATGCGAGATATCCGGAATTA 1756
Db 1680 AGTGAAGAGAAACTGTGAGATTAACTTGAAGATATGCGAGATATCCGGAATTA 1739
QY 1757 CCTCAAGAACTTTTGAAGAAATCACCACCAACCAATACACTCACTCACTCAAAAGTACAT 1816
Db 1740 CCTCAAGAACTTTTGAAGAAATCACCACCAACCAATACACTCACTCACTCAAAAGTACAT 1799
QY 1817 GGAATCTTATTTGGGCTTGAAGAAACGAGATGAAGCACTTCTCAACCTGCTTCTCAATAG 1876
Db 1800 GGAATCTTATTTGGGCTTGAAGAAACGAGATGAAGCACTTCTCAACCTGCTTCTCAATAG 1859
QY 1877 GGAATGCTTTCAGAGCAAGTCTCTCACTGTGTGTACACAGCAGAGCACTGATCAGTC 1936
Db 1860 GGAATGCTTTCAGAGCAAGTCTCTCACTGTGTGTGTACACAGCAGAGCACTGATCAGTC 1919
QY 1937 ACAGCATACAGCTGTCCACTGAAAGAACTGTCTCAACAGCTGAATCAATTTGTT 1996
Db 1920 ACAGCATACAGCTGTCCACTGAAAGAACTGTCTCAACAGCTGAATCAATTTGTT 1979
|||||

QY	1997	AGCTTAATAGATAAAATCCAGACTTACTTACGCTTTATATCCCTTTATTATCATTAATAAC	2056
Db	1980	AGCTTAATAGATAAAATCCAGACTTACTTACGCTTTATATCCCTTTATTATCATTAATAAC	2039
QY	2057	TGTGAACCTGACGACGACCAATGGAAACATTATTACTGAGCTCGAGATTCGAGCTCGG	2116
Db	2040	TGTGAACCTGACGACGACCAATGGAAACATTATTACTGAGCTCGAGATTCGAGCTCGG	2099
QY	2117	AACCTTAGTCTTATCTGATCCAGACAGACGACACCTTAGTATCTGCCAACTAATG	2176
Db	2100	AACTTAGTCTTATCTGATCCAGACAGACGACACCTTAGTATCTGCCAACTAATG	2155
QY	2177	AGTTAATATAA	2187
Db	2160	AGTTAATATAA	2170
RESULT 5			
LOCUS	AY357925	2187 bp	mRNA
DEFINITION	Macaca faecicularis beestrophin (VMD2) mRNA, complete cds.		
ACCESSION	AY357925		
VERSION	AY357925.1	GI:34013782	
KEYWORDS			
SOURCE	Macaca faecicularis (crab-eating macaque)		
ORGANISM	Macaca faecicularis		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecidae; Macaca.		
AUTHORS	1 (bases 1 to 2187) Okamoto,H., Umeda,S., Suzuki,M.T., Yoshihawa,Y., Tanaka,Y. and Iwata,Y.		
TITLE	Molecular Cloning of VMD2 Gene from Cynomolgus Monkey (Macaca faecicularis)		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 2187)		
AUTHORS	Okamoto,H., Umeda,S., Suzuki,M.T., Yoshihawa,Y., Tanaka,Y. and Iwata,Y.		
TITLE	Direct Submission		
JOURNAL	Submitted (01-AUG-2003) National Tokyo Medical Center, National Institute of Sensory Organs, 2-5-1, Higeashioka, Meguro-ku, Tokyo 152-0021, Japan		
FEATURES			
source	Location/Qualifiers 1..2187 /organism="Macaca faecicularis" /mol_type="mRNA" /db_xref="taxon:9541" 1..2187 /gene="VMD2" 67..1824 /gene="VMD2" /codon_start=1 /product="beestrophin" /protein_id="AA056049.1" /db_xref="GI:34013783"		
gene	translating="MTTITTSQVANAIGSTSRLLCMKRGITIKLIGEPITPLCYYY IIRIRLALFVEEQLMKEKTLKCDSTYQILPISFVLGFTVTLVTRMWOYNLPLP IDRLMSLVGVEBQDEQGRLLRRLIRANIGVULIRSVTVYKRFPSOHLVOVA GPMPLPAEHKQLEKLSLPHNMFVPMWPMASMKAMIGGRJRDPILOSLLNNNTL TQGHLLAYDMVISIPLYVTQVTAAYVSFPLTCLVGRQFLPAKAYRGEHLIVVPR TFLOFFPYGMLKYAEOLINFGEDDDPFRTMTVDRLQVSLAIDEMODLPRMER DMYNEPPEPPTPTAASQPRASMGSTFNISLNKEMEYQPOBEDKDAHTGITIG FLGQSDHHPPGANSRTKLMPKRESLLHSGLEKPNKAVIQNVRGLEDNAMKLKAW DAFKSAPLYQRPGYSAQTPLSPTPMFPPEPSVLSGLHSIVTGIQDKSLKVSIGS AKNSSELISGDMERBEVSHVARKTVEFNLTDMPETPINHLKEPLELSTYIMATW LKDHVDPYWALENRDEAHS"		
CDS			
ORIGIN			
Query Match	90.7%	Score 2022.6	DB 9
Best Local Similarity	95.7%	Pred. No. 0	Length 2187
Matches 2092	Conservative	0	Mismatches 89
			Indels 6
			Gaps 1

QY	39	GGATTCATCGAACCAACCTGGAAACCCACCTGACCCAGGCCACCTGCTGACGCCACCTGC	98
Db	1	GGATTCACCGGATCCACCTGGAAACCCACCTGACCCAGGCCACCTGCTGACGCCAAAGTGC	60
QY	99	CTGGCCATGACCATCATCTTACACAAAGCCAAATGGCTAAATGCCCCGTTAGCTCTCTTCC	158
Db	61	CTGGCCATGACCATCATCTTACACAGGCCAAATGGCTAAATGCCGTTAGCTCTCTTCC	120
QY	159	CGCTGCTGCTGTGTGTGGGGGGGAGCATCTACAAGCTGCTATATATGCGAGTTCTTAAATC	218
Db	121	CGCTGCTGCTGTGTGTGGGGGGGAGCATCTACAAGCTGCTATATATGCGAGTTCTTAAATC	180
QY	219	TTCTGCTCTGCTACTACTATATCCGCTTTATTTATAGCTGGCCCTTACGGAAGAACAA	278
Db	181	TTCTGCTCTGCTACTACTATATCCGCTTTATTTATAGCTGGCCCTTACGGAAGAACAA	240
QY	279	CAGCTGATGTTTGAAACGTACCTGTATTTGCGACAGCTACATCCAGCTCATCCCATTT	338
Db	241	CAGCTGATGTTTGAAACGTACCTGTATTTGCGACAGCTACATCCAGCTCATCCCATTT	300
QY	339	TCCTTCTGCTGAGGCTTCTACGTACCGCTGTCTGTGACCCGCTGGTGGACACAGTACAG	398
Db	301	TCCTTCTGCTGAGGCTTCTACGTACCGCTGTCTGTGACCCGCTGGTGGAAACAAGTACAG	360
QY	399	AACCTGCCGTGGGCCGAGCCGCTCATGAGCCTGTGTCCGCTTCTGTGAAGCGAAGAC	458
Db	361	AACCTGCCGTGGGCCGAGCCGCTCATGAGCTTGTGTGTCCGCTTCTGTGAAGCGAAGAC	420
QY	459	GAGCAAGGCCGGCTGTGCTGGGGGGGACGCTATCCGCTACGGCAACCTGGGGAAGCTGCTC	518
Db	421	GAGCAAGGCCGGCTGTGCTGGGGGGGACGCTATCCGCTACGGCAACCTGGGGAAGCTGCTC	480
QY	519	ATCTGTGCGACAGCTGTACAGCACCGCAGTCTACAAGGCTTCCCCAGGCCACAGACCTGTG	578
Db	481	ATCTGTGCGACAGCTGTACAGCACCGCAGTCTACAAGGCTTCCCCAGGCCACAGACCTGTG	540
QY	579	CAAGCAGGCTTATATGACTCCGCGCAGAACACAGACGTTGGAGAACTGAGCCTTACACAC	638
Db	541	CAAGCAGGCTTATATGACCCCGCGCAGAACACAGACGTTGGAGAACTGAGCCTTACACAC	600
QY	639	AACTATGTTTGGGGTCCCTGGGTGTGTTTCCCACTGTACATGAAGGCGTGGCTTGGGA	698
Db	601	AACTATGTTTGGGGTCCCTGGGTGTGTTTCCCACTGTACATGAAGGCGTGGCTTGGGA	660
QY	699	GGTGGAATCCGGGACCCCTATCTGTCTCAGAGCCTGTGAACGAGATGAACACTTGGGT	758
Db	661	GGTGGAATCCGGGACCCCTATCTGTCTCAGAGCCTGTGTGAACGAGATGAACACTTGGGT	720
QY	759	ACTCAGTGTGACACTGTATGCTTACGACCTGGAATTAGTATCCGACTGTTATACACAG	818
Db	721	ACTCAGTGTGACACTGTATGCTTACGACCTGGAATTAGTATCCGACTGTTATACACAG	780
QY	819	GTGTGACTGTGGGGGTGTACAGCTTCTTCTGACTTGTCTAGTTGGGCGGCAGTTTCTG	878
Db	781	GTGTGACTGTGGGGGTGTACAGCTTCTTCTGACTTGTCTAGTTGGGCGGCAGTTTCTG	840
QY	879	AAACCAGCAAGGCTTACCTGTGGCATGAGCTGGAACCTGTTGTGGCCGCTTCAACGTTT	938
Db	841	AAACCAGCAAGGCTTACCTGTGGCATGAGCTGGAACCTGTTGTGGCCGCTTCAACATTC	900
QY	939	CTGCAGTTCTTCTTATATGTTGCTGTGCTGAAGGTGGCAGACGCTCATCAACCCCTTT	998
Db	901	CTGCAGTTCTTCTTATATGTTGCTGTGCTGAAGGTGGCAGACGCTCATCAACCCGTTT	960
QY	999	GGAGAGATGATGATGATTTTGAACCACTGGAATTGTCCAGAGAAATTTGCAGGTGTCC	1058
Db	961	GGAGAGAGATGATGATTTTGAACCACTGGAATTGTCCAGAGAAATTTGCAGGTGTCC	1020
QY	1059	CTGTGTGCTGTGATGATGATGACACAGGACCTGTGCTGGATGAGGCGGGAACATGATCTGG	1118
Db	1021	CTGTGTGCTGTGATGATGATGACACAGGACCTGTGCTGGATGAGGCGGGAACATGATCTGG	1080
QY	1119	AATTAAGCCGAGCCACAGCCCCCTTACACAGCTGTTCCGCCAGTTCCGTGAGCCTCC	1178

```

Db      1081 AATGAGCTGAGCAGACCCCTTACAGAGCTCTCCGCCAGATTCGATCCAGCTCTCC 1140
Qy      1179 TTTATGGGCTCCACCTTCAACATGAGCTGAACAAAGAGAGATGAGATTCAGCCCAAT 1238
Db      1141 TTTATGGGCTCCACCTTCAACATGAGCTGAACAAAGAGAGATGAGATTCAGCCCAAT 1200
Qy      1239 CAGAGAGACGAGAGAGATGCTCAGCTGAGCATTTGGCGCTTCCAGGCTTCAGGCTTC 1298
Db      1201 CAGAGAGACAAAGAGAGAGCTCAGCTGAGCATTTGGCGCTTCCAGGCTTCAGGCTTC 1260
Qy      1299 CAGATATCAATCTCTCCAGAGGCAAACTCAAGAGCAAACTACTGTGCGCCAGAGAGAA 1358
Db      1261 CAGATATCAATCTCTCCAGAGGCAAACTCAAGAGCAAACTACTGTGCGCCAGAGAGAA 1320
Qy      1359 TCCCTTCTCCAGAGGCGCTGCCCAAAAACCAAGAGCGCCCAACAGAACTTTAGGGGC 1418
Db      1321 TCCCTTCTCCAGAGGCGCTGCCCAAAAACCAAGAGCGCGCATCAAGAACTTTAGGGGC 1380
Qy      1419 CAGAGAGACAAAGAGCTGGAAGCTTAAAGCTGTGAGAGCGCTTCAAGCTGTGCGCACTG 1478
Db      1381 CTGGAAGACAAAGAGCTGGAAGCTTAAAGCTGTGAGAGCGCTTCAAGCTGTGCGCACTG 1440
Qy      1479 TATCAGAGGCGAGGCTACTCAAGTCCCCAAGAGCGCTTCAAGCTTCCATGCTTC 1538
Db      1441 TATCAGAGGCGAGGCTACTCAAGTCCCCAAGAGCGCTTCAAGCTTCCATGCTTC 1500
Qy      1539 TTTCCCTTGAAGACATCAGAGCGCGCTCAAGCTTCAAGTTCACAGGATGAGGATGAGCA 1598
Db      1501 TTTCCCTTGAAGACATCAGAGCGCGCTCAAGCTTCAAGTTCACAGGATGAGGATGAGCA 1560
Qy      1599 GACAAAGCTTAAAGCTGTGAGTTCGGGCGCAAGAAAGTTTGAATTGCTCTCAGAG 1658
Db      1561 GACAAAGCTTAAAGCTGTGAGTTCGGGCGCAAGAAAGTTTGAATTGCTCTCAGAG 1620
Qy      1659 AGCGATGGGCGCTTGAAGAGACCCAGAGATCCCGGAAATCACTCAAGAACTTTGGAACAA 1778
Db      1621 AGCGATGGGCGCTTGAAGAGACCCAGAGATCCCGGAAATCACTCAAGAACTTTGGAACAA 1740
Qy      1719 TTTAACTGAGCGATATGCGAGAGATCCCGGAAATCACTCAAGAACTTTGGAACAA 1778
Db      1681 TTTAACTGAGCGATATGCGAGAGATCCCGGAAATCACTCAAGAACTTTGGAACAA 1740
Qy      1779 TCAGCAACCAACATCACTCACTCACTCAAGATCACTCAAGATCTTATTTGGGCTTGGAA 1838
Db      1741 TCAGCAACCAACATCACTCACTCACTCAAGATCACTCAAGATCTTATTTGGGCTTGGAA 1800
Qy      1839 AACAGGATGGAAGCATTTCTTAAGCTGCTTCTTAATGAGGATGCTTCCAGGCTTGGAA 1898
Db      1801 AACAGGATGGAAGCATTTCTTAAGCTGCTTCTTAATGAGGATGCTTCCAGGCTTGGAA 1860
Qy      1899 CTCACCTGTGTATGACAGAGAGAGACTGATCCAGCTCAAGCATCAAGCTGTCCACAC 1958
Db      1861 CTCACCTGTGTATGACAGAGAGAGACTGATCCAGCTCAAGCATCAAGCTGTCCACAC 1920
Qy      1959 TGAAGAAAGTGTCTCAACAGAGCTGAAATCAAGGTTAGCTTAATAGATTA-----A 2012
Db      1921 TGAAGAAAGTGTCTCAACAGAGCTGAAATCAAGGTTAGCTTAATAGATTAATCA 1980
Qy      2013 ATCCAGATCACTTCAAGCTTAAATGCTTTTATTCATTAATAAAGCTGGAAGCTAGAGCTG 2072
Db      1981 GTCCAGAGATCACTTCAAGCTTAAATGCTTTTATTCATTAATAAAGCTGGAAGCTAGAGCTG 2040
Qy      2073 AACGATTTGGAAGCACTTTATCTCAAGCTGTGATTCAGAGTGGGAACTTTAGCTTATC 2132
Db      2041 AACGATTTGGAAGCACTTTATCTCAAGCTGTGATTCAGAGTGGGAACTTTAGCTTATC 2100
Qy      2133 TGAATCCAGAGAGCGCACCTTATGATATGCGCAAACTAATGAGTTTAATAAAGCA 2192
Db      2101 TGAATCCAGAGAGCGCACCTTATGATATGCGCAAACTAATGAGTTTAATAAAGCA 2160
Qy      2193 ATACTCTGTTAAAAAAGGCTTTAATTAATAAAGCA 2219

```

```

Db      2161 ATACTCTGTTAAAAAAGGCTTTAATTAATAAAGCA 2187

RESULT 6
BD136721
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
OS Homo sapiens (human)
PN JP 2002504559-A/3
PD 12-FEB-2002
PF 22-FEB-1999 JP 2000533447
PR 25-FEB-1998 US 60/075941; 18-DEC-1998 US 60/112926
PT KONSTANTIN PETRUKHIN, THOMAS C CASKEY, MICHAEL METZKER, CLAES PI
WADIELIUS
PC C07K16/18, C07K14/47, C12N5/10, C12N15/09, C12P19/34, C1201/68 // PC
C12P21/08,
PC C12N5/00, C12N15/00
CC Best's macular dystrophy gene
FT Key Location/Qualifiers
FT source 1. 2429 Location/Qualifiers
FT PT /organism='Homo sapiens (human)'.
FEATURES
source
1. 2429 Location/Qualifiers
/organism='Homo sapiens'
/mol_type='genomic DNA'
/db_xref='taxon:9606'
ORIGIN
Query Match 90.3%; Score 2013; DB 6; Length 2429;
Best Local Similarity 91.6%; Pred. No. 0;
Matches 2226; Conservative 0; Mismatches 0; Indels 203; Gaps 1;
Qy      1 CAGGAGTCCCAACGAGCTGATGCGCAGACCTTTCTGTGATCATCGAACCTGAGAA 60
Db      1 CAGGAGTCCCAACGAGCTGATGCGCAGACCTTTCTGTGATCATCGAACCTGAGAA 60
Qy      61 CCCCACTTGAACCAAGCCCACTGCTGAGAGCCCACTGCTGCGCATGACCACTTACA 120
Db      61 CCCCACTTGAACCAAGCCCACTGCTGAGAGCCCACTGCTGCGCATGACCACTTACA 120
Qy      121 CAAAGCAAGTGGCTTAATGCGCGCTTAAAGCTCTTCTCCGCTGCTGTGCTGGCGGG 180
Db      121 CAAAGCAAGTGGCTTAATGCGCGCTTAAAGCTCTTCTCCGCTGCTGTGCTGGCGGG 180
Qy      121 CAAAGCAAGTGGCTTAATGCGCGCTTAAAGCTCTTCTCCGCTGCTGTGCTGGCGGG 180
Db      121 CAAAGCAAGTGGCTTAATGCGCGCTTAAAGCTCTTCTCCGCTGCTGTGCTGGCGGG 180
Qy      181 GCAGCATCTACAAGCTGCTATATGAGCGAGTCTTAAATGCTGCTGCTCACTACATCA 240
Db      181 GCAGCATCTACAAGCTGCTATATGAGCGAGTCTTAAATGCTGCTGCTCACTACATCA 240
Qy      241 TCCGCTTATTTAATAGAGTGGCCCTCAAGAGAAACAACAGCTGATGTTGAAGAACTGA 300
Db      241 TCCGCTTATTTAATAGAGTGGCCCTCAAGAGAAACAACAGCTGATGTTGAAGAACTGA 300
Qy      301 CTGTGTATTTGGAAGAGCTACATCCAGCTCATCCCATTTTCTGTGCTGGGCTTCAAG 360
Db      301 CTGTGTATTTGGAAGAGCTACATCCAGCTCATCCCATTTTCTGTGCTGGGCTTCAAG 360
Qy      361 TGAGCGTGTGCTGACCCGCTGTGGAACAGTACAGAAAGCTGCGTGGCGGACCGCC 420
Db      361 TGAGCGTGTGCTGACCCGCTGTGGAACAGTACAGAAAGCTGCGTGGCGGACCGCC 420
Qy      421 TCATGAGCTGTGTGCGGCTTCTGTCGAAGGCAAGAGAGCAAGAGCGGCTGCGGCG 480

```

Db 421 TCATAGAGCTGGTGTGCGGGCTTGTCGAAGGCAAGAGCAAGGCCGCTGTCGGGC 480
Qy 481 GCAAGCTCATCCGCTACGCGCAACCTGGGCAAGTGTCTCATCTTGGCGACACCTGACACCG 540
Db 481 GCAAGCTCATCCGCTACGCGCAACCTGGGCAAGTGTCTCATCTTGGCGACACCTGACACCG 540
Qy 541 CAGTCTACAGAGCGCTTCCCGACGCGCCAGCACTGTGTGCAAGCAGGCTTATGATCCGG 600
Db 541 CAGTCTACAGAGCGCTTCCCGACGCGCCAGCACTGTGTGCAAGCAGGCTTATGATCCGG 600
Qy 601 CAGAACACAGAGCGTGGAGAACTGAGCTTACCAACAATGTTCTTGGGCTGAGG 660
Db 601 CAGAACACAGAGCGTGGAGAACTGAGCTTACCAACAATGTTCTTGGGCTGAGG 660
Qy 661 TGTGGTGGCCCACTGTGCAATGAAGGCGTGGCTTGGAGTCCAGGACCTATCC 720
Db 661 TGTGGTGGCCCACTGTGCAATGAAGGCGTGGCTTGGAGTCCAGGACCTATCC 720
Qy 721 TGCTCCAGAGCGCTGCTGAACGAGATGAACACCTTGCTGCTGATGCTGATG 780
Db 721 TGCTCCAGAGCGCTGCTGAACGAGATGAACACCTTGCTGCTGATGCTGATG 780
Qy 781 CCTACGACCTGATAGTATCCCACTGTGTATACAGAGTGTGACTGTGGCGGTATCA 840
Db 781 CCTACGACCTGATAGTATCCCACTGTGTATACAGAGTGTGACTGTGGCGGTATCA 840
Qy 841 GCTTCTTCTCTGACTTGTCTAGTGGGCGGCAAGTCTTGAACCCAGCCAGGCTTACCTG 900
Db 841 GCTTCTTCTCTGACTTGTCTAGTGGGCGGCAAGTCTTGAACCCAGCCAGGCTTACCTG 900
Qy 901 GCCATGAGCTGGAACCTGCTGTGCGCGTCTTCAAGTCTTCTGATGTTG 960
Db 901 GCCATGAGCTGGAACCTGCTGTGCGCGTCTTCAAGTCTTCTGATGTTG 960
Qy 961 GCTGCTGTA----- 970
Db 961 GCTGCTGTAAGGTTGGGCTCTCCAGGGCCCTGCTGGGCTGAGGATGGCCAGAGGGGTC 1020
Qy 971 ----- 970
Db 1021 ATGGCCAGAGCTGCTTGAGACGAGATGCAAGTGTCAAGAAAGAAAGTCTCACGGGTAG 1080
Qy 971 ----- 970
Db 1081 AAAGCAGCCAGGCGTGTGGCGCACACTGTATCCAGCTACTCGGAGGCTGAGGCGAG 1140
Qy 971 -----GGTGGCAGAGCAGCTCATCAACCCCTT 997
Db 1141 GAGAAATGCTTGAAACCCGGGAGCGGAGGTTGTGTGGCAGAGCAGCTCATCAACCCCTT 1200
Qy 998 TGGAGAGATGATGATGATTTTGGAGCAACAATGATGTCGACAGAAATTTGCAAGTGTG 1057
Db 1201 TGGAGAGATGATGATGATTTTGGAGCAACAATGATGTCGACAGAAATTTGCAAGTGTG 1260
Qy 1058 CCTGTGGCTGTGATGATGATGACACAGAACTGTGCTGAGATGAGCGAGCAATGATCTG 1117
Db 1261 CCTGTGGCTGTGATGATGATGACACAGAACTGTGCTGAGATGAGCGAGCAATGATCTG 1320
Qy 1118 GAATTAAGCCGAGCAGACAGCCCTTACACAGCTCTTCCGCCAGTTCGGTCAGGCTTC 1177
Db 1321 GAATTAAGCCGAGCAGACAGCCCTTACACAGCTCTTCCGCCAGTTCGGTCAGGCTTC 1380
Qy 1178 CTTTATGGGCTTCCACTTCAACATCAGCTGTAACAAAGAGAGATGAGTTCAGCCCAA 1237
Db 1381 CTTTATGGGCTTCCACTTCAACATCAGCTGTAACAAAGAGAGATGAGTTCAGCCCAA 1440
Qy 1238 TCAGAGAGACGAGAGATGCTCAGGCTGAGATCAATTTGGCGCTTCTTAAGGCTTCAATC 1297
Db 1441 TCAGAGAGACGAGAGATGCTCAGGCTGAGATCAATTTGGCGCTTCTTAAGGCTTCAATC 1500
Qy 1298 CCATGATCAGCATCTCCAGGCGAACTCAAGAGCAAACTAGTGGGCCAAAGAGGA 1357
Db 1357 CCATGATCAGCATCTCCAGGCGAACTCAAGAGCAAACTAGTGGGCCAAAGAGGA 1357

Db 1501 CCATGATCAGCATCTCCAGGCGAACTCAAGAGCAAACTAGTGGGCCAAAGAGGA 1560
Qy 1358 ATCCCTTTCACAGAGGCGCTGCCAAACCAAGGCAAGCAACGATTAAGGG 1417
Db 1561 ATCCCTTTCACAGAGGCGCTGCCAAACCAAGGCAAGCAACGATTAAGGG 1620
Qy 1418 CCAGAGACAAACAAGGCTGGAAGCTTAAGGCTGTGACGCGCTTCAAGTCTGCGCCACT 1477
Db 1621 CCAGAGACAAACAAGGCTGGAAGCTTAAGGCTGTGACGCGCTTCAAGTCTGCGCCACT 1680
Qy 1478 GTATCAGAGGCGAGCTTACTACAGTGGCCCAAGAGCGCCCTCAGCCCACTCCATGTT 1537
Db 1681 GTATCAGAGGCGAGCTTACTACAGTGGCCCAAGAGCGCCCTCAGCCCACTCCATGTT 1740
Qy 1538 CTTCCCTTAAACCATCAGAGGCGCTCAAGCTTCAAGTGTGACAGGCAATGAGACCA 1597
Db 1741 CTTCCCTTAAACCATCAGAGGCGCTCAAGCTTCAAGTGTGACAGGCAATGAGACCA 1800
Qy 1598 AGACAAAGCTTAAAGCTGTGAGTTCGGGCGCAAGAAAGTTTGAATTGCTCTCAGA 1657
Db 1801 AGACAAAGCTTAAAGCTGTGAGTTCGGGCGCAAGAAAGTTTGAATTGCTCTCAGA 1860
Qy 1658 GAGCGATGGGCGCTTGTATGAGACACCGAAGATCTCAAGTGAAGAGAAAATGTGGA 1717
Db 1861 GAGCGATGGGCGCTTGTATGAGACACCGAAGATCTCAAGTGAAGAGAAAATGTGGA 1920
Qy 1718 GTTTAACTGACGATATGCGAGAGATCCCGGAAATCACTCAAAAGAACTTTGGAA 1777
Db 1921 GTTTAACTGACGATATGCGAGAGATCCCGGAAATCACTCAAAAGAACTTTGGAA 1980
Qy 1778 ATCAACCAACCAATACACACTACCTCAAGATCACTGATCTTATGGGCGCTTGA 1837
Db 1981 ATCAACCAACCAATACACACTACCTCAAGATCACTGATCTTATGGGCGCTTGA 2040
Qy 1838 AACAGGATGAAGCACTTCTTAACTGCTTCTTATGAGGATGCTTGGCGCAGGT 1897
Db 2041 AACAGGATGAAGCACTTCTTAACTGCTTCTTATGAGGATGCTTGGCGCAGGT 2100
Qy 1898 CCTACCTGTGTATACACAGAGACACTGATCCAGTCAACAGCATACAGTGTCCACA 1957
Db 2101 CCTACCTGTGTATACACAGAGACACTGATCCAGTCAACAGCATACAGTGTCCACA 2160
Qy 1958 CTGAAGAAAGTGTCTTACACAGAGCTGATCAAAATGTTAGCTTATGATTAATATCC 2017
Db 2161 CTGAAGAAAGTGTCTTACACAGAGCTGATCAAAATGTTAGCTTATGATTAATATCC 2220
Qy 2018 AGACTACTTCAAGCTTATATGCTTATGATTAATATGATTAATATGATTAATATCC 2077
Db 2221 AGACTACTTCAAGCTTATATGCTTATGATTAATATGATTAATATGATTAATATCC 2280
Qy 2078 TTGAAACATTTAACTCAGACTCTGGAATTCAGAGTTCGGAAACCTTATGTTCTGAAT 2137
Db 2281 TTGAAACATTTAACTCAGACTCTGGAATTCAGAGTTCGGAAACCTTATGTTCTGAAT 2340
Qy 2138 CCAAGACAGCAACCTTATGATTAAGTGTGCAAGCTTATGATTAATATGATTAATATCT 2197
Db 2341 CCAAGACAGCAACCTTATGATTAAGTGTGCAAGCTTATGATTAATATGATTAATATCT 2400
Qy 2198 GCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2226
Db 2401 GCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2429

RESULT 7
AF057170 2420 bp mRNA linear PRI 17-OCT-1998
LOCUS Homo sapiens beactrophin (VMD2) mRNA, alternatively spliced product,
DEFINITION complete cds.
ACCESSION AF057170
VERSION AF057170.1 GI:3335160
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens


```

QY 1298 CCATGATCACCATCTCTCCAGAGGCAAACTCAAGACCCAACTACTGTGGCCCAAGAGGA 1357
DB 1501 CCAATGATCACCATCTCTCCAGAGGCAAACTCAAGACCCAACTACTGTGGCCCAAGAGGA 1560
QY 1358 ATCCCTTCTCCAGAGGCGCTGCCCAAAACCAAGAGGAGCCCAACGAACTTTAGGGG 1417
DB 1561 ATCCCTTCTCCAGAGGCGCTGCCCAAAACCAAGAGGAGCCCAACGAACTTTAGGGG 1620
QY 1418 CCAGGAAGACAAACAAGGCGTGAAGCTTAAAGCTGTGAGCCCTTCAAGCTTGCCCACT 1477
DB 1621 CCAGGAAGACAAACAAGGCGTGAAGCTTAAAGCTGTGAGCCCTTCAAGCTTGCCCACT 1660
QY 1478 GTATCAGAGGCGCAGGCTACTACAGTCCCAAGAGCCCTCAGAGCCCACTCCCATGTT 1537
DB 1681 GTATCAGAGGCGCAGGCTACTACAGTCCCAAGAGCCCTCAGAGCCCACTCCCATGTT 1740
QY 1538 CTTCCCTTACAGAACATCAGCGCGCTCAAGCTTCAAGTGTCAAGGCTATGACACCA 1597
DB 1741 CTTCCCTTACAGAACATCAGCGCGCTCAAGCTTCAAGTGTCAAGGCTATGACACCA 1800
QY 1598 AGACAAAAGCTTAAAGACTGTGAGTTCTGGGGCCCAAGAAAGTTTGAATTGCTCTAGA 1657
DB 1801 AGACAAAAGCTTAAAGACTGTGAGTTCTGGGGCCCAAGAAAGTTTGAATTGCTCTAGA 1860
QY 1658 GAGCGATGGGGCTTGTGATGAGCACCCAAGATATCTCAAGTGAAGAGGAAACCTGAGA 1717
DB 1861 GAGCGATGGGGCTTGTGATGAGCACCCAAGATATCTCAAGTGAAGAGGAAACCTGAGA 1920
QY 1718 GTTTTACCTGACGGATATGCCAGAGATCCCGAAATCACCCTCAAGAACCTTTGAAACA 1777
DB 1921 GTTTTACCTGACGGATATGCCAGAGATCCCGAAATCACCCTCAAGAACCTTTGAAACA 1980
QY 1778 ATCCCAACCAACTATACACTACACTCAAGATCAAGATCACTGATTTGGGCTTGA 1837
DB 1981 ATCCCAACCAACTATACACTACACTCAAGATCAAGATCACTGATTTGGGCTTGA 2040
QY 1838 AATCAGGAGATGAGACATCTCTAAGCTGCTTCTTAATGGGAGATGCTTCCGACAGCAGT 1897
DB 2041 AATCAGGAGATGAGACATCTCTAAGCTGCTTCTTAATGGGAGATGCTTCCGACAGCAGT 2100
QY 1898 CCTCAGCTGTGTGTACACAGACAGAGACACTGATCCAGTCAAGCCATTCAGCTGTACA 1957
DB 2101 CCTCAGCTGTGTGTACACAGACAGAGACACTGATCCAGTCAAGCCATTCAGCTGTACA 2160
QY 1958 CTGAAGAACTGTCTCTACACAGCCTGATTAATGGTTAGCTTAAATGATTAATAATCCC 2017
DB 2161 CTGAAGAACTGTCTCTACACAGCCTGATTAATGGTTAGCTTAAATGATTAATAATCCC 2220
QY 2018 AGACTACTTCAAGCTTAAATGCTTTTAAATGATTAATAATACTGAAAGCTTACGAAACA 2077
DB 2221 AGACTACTTCAAGCTTAAATGCTTTTAAATGATTAATAATACTGAAAGCTTACGAAACA 2280
QY 2078 TTGGAACATTTAATCTAGACTCTGAGATTCAGAGTCCGGAACCTTATGTTTATCTGAT 2137
DB 2281 TTGGAACATTTAATCTAGACTCTGAGATTCAGAGTCCGGAACCTTATGTTTATCTGAT 2340
QY 2138 CCAAGACAGCCACCTTATGATTAATGCTGCAAACTTAATGATTAATAATAAATCT 2197
DB 2341 CCAAGACAGCCACCTTATGATTAATGCTGCAAACTTAATGATTAATAATAAATCT 2400
QY 2198 CGTTAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2217
DB 2401 CGTTAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2420

```

```

RESULT 8
BC015220 2441 bp mRNA linear PRI 07-OCT-2003
LOCUS BC015220
DEFINITION Homo sapiens vitelliform macular dystrophy (Best disease,
bestrophin), mRNA (cDNA clone MGC:17761 IMAGE:3877806), complete
cds.
ACCESSION BC015220

```

```

VERSION BC015220.2 GI:34783515
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2441)
AUTHORS Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Hopkins, R.F., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Uedlin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loggellano, N.A., Peters, G.J.,
Abramson, R.D., Muliahy, S.J., Boeak, S.A., McEwan, P.J.,
McKernan, K.U., Malek, U.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Huily, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whitton, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalins, D.E.,
Scherer, A., Schein, J.E., Jones, S.J., and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
JOURNAL MEDLINE
12477932
PUBMED
REFERENCE 2 (bases 1 to 2441)
AUTHORS Strausberg, R.
TITLE Direct Submission
Submitted (01-OCT-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
On Sep 16, 2003 this sequence version replaced gi:2195361.
REMARK COMMENT
Contract: MGC help desk
Email: gcgds-remail.nih.gov
Tissue Procurement: DCTD/DRP/Gazdar
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LUNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www.sngc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LUNL at: http://image.llnl.gov
Series: IRAX Plate: 14 Row: b Column: 9.
FEATURES
source
1..2441
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:17761 IMAGE:3877806"
/tissue_type="Lung, large cell carcinoma"
/clone_id="NIH MGC_68"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
1..2441
/gene="VMD2"
/notes="synonyms: BMD, BEST"
/db_xref="locusid:7439"
/db_xref="MIM:153700"
573..2069
/gene="VMD2"
/codon_start=1
/product="VMD2 protein"

```

/protein_id="AAH15220.1"
/db_xref="GI:22800661"
/db_xref="LOCUSID:7439"
/db_xref="MIM:153700"
/translation="MFEKLTLYCDSYIQLIPISFVLGFVYTLVTRMNOYENLPMD
RLMSLVSGFVBKDEOGLRLRLTLYRANLGNVLISVSVAAYKPPSPKHVOAGP
MTPAEHKLQELSLPHNMFVWVWPNLMSKMLGGRIIDILLQSLNEMTLTQ
CGHLYADWISIPLYOVTVAVYSPFLICVIGROPTINRAKYPGHELDLVPTFL
LOPFYVGMILKVLSLAVDEMNDLPMEPPMYNKKRPPPTPAASQKFRASFWEST
FNISLNKEBMEFOPNDEEDNAHGIIGRLGLOSIDHHPRANSSRTKLMPKRESIL
HEGLPKNHKCAKONVROGEBNKMLKVAIVAFKSAPLYORGYSAPOPLSPWPF
PLPSASKHSHTVIGIDTKOKSLTKVSSAKSKSPFLISBEGALMEHPVSGVRRKTV
EFNLIDMPEIPEHNHKEPLKOSPLNIHTLIKDHMDPYALENDRBAS"
576..1427
misc_feature
/gene="VMD2"
/note="DUF289; Region: Putative membrane protein. This
family called family 8 in, may be a transmembrane protein
The specific function of this protein is unknown"
/db_xref="CDL:pfam01062"

Query Match 80.4%; Score 1793; DB 9; Length 2441;
Best Local Similarity 95.6%; Pred. No. 0;
Matches 1887; Conservative 0; Mismatches 5; Indels 81; Gaps 1;

QY 257 GCTGGCCCTCAACGAGAACACACAGCTGATGTTGAGAACTGACTGTGATTGGACAG 316
DB 545 GCTGGCCCTCAACGAGAACACACAGCTGATGTTGAGAACTGACTGTGATTGGACAG 604
QY 317 CTACATCAGCTCATCCCATTTCCCTTCGTCGCGGCTTCTACAGTACGCTGCTGTGAC 376
DB 605 CTACATCAGCTCATCCCATTTCCCTTCGTCGCGGCTTCTACAGTACGCTGCTGTGAC 664
QY 377 CGCGTGTGAAACAGTACGAGAACCTGCGTGCCTGACCGACCGCTCATGAGGCTGTGTC 436
DB 665 CGCGTGTGAAACAGTACGAGAACCTGCGTGCCTGACCGACCGCTCATGAGGCTGTGTC 724
QY 437 GGGCTTCGTGGAAGGAAGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 496
DB 725 GGGCTTCGTGGAAGGAAGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 784
QY 497 CGCCAACTGGGCAACGTGCTCATCTGCGCAGCTGACAGCCGAGCTACAGCGCTT 556
DB 785 CGCCAACTGGGCAACGTGCTCATCTGCGCAGCTGACAGCCGAGCTACAGCGCTT 844
QY 557 CCCGAGCGCCAGCAGCTGTGTCAGAGAGGCTTTATGACTCGGCAAGAACAGCAGTT 616
DB 845 CCCGAGCGCCAGCAGCTGTGTCAGAGAGGCTTTATGACTCGGCAAGAACAGCAGTT 904
QY 617 GAGAGAACTGAGCCCAACACACATGTTCTGGGTGCGCGGTGCTGTTGGCAACT 676
DB 905 GAGAGAACTGAGCCCAACACACATGTTCTGGGTGCGCGGTGCTGTTGGCAACT 964
QY 677 GTCAATGAAGCGTGTGAGGTGCAATCGGAGCCCTATCTGTGTCAGAGCGTGT 736
DB 965 GTCAATGAAGCGTGTGAGGTGCAATCGGAGCCCTATCTGTGTCAGAGCGTGT 1024
QY 737 GAAAGAGTGAACACCTTGCTACTCACTGTGTCACCTGTATGCTTCACTGATGAG 796
DB 1025 GAAAGAGTGAACACCTTGCTACTCACTGTGTCACCTGTATGCTTCACTGATGAG 1084
QY 797 TATCCCACTGTGTATACACAGGTGTGATGTCGAGTGTAGAGCTTCTTCTGACTTG 856
DB 1085 TATCCCACTGTGTATACACAGGTGTGATGTCGAGTGTAGAGCTTCTTCTGACTTG 1144
QY 857 TCTAGTTGGGCGGAGGTTTCTGAACCTGAGCAAGCTTACCTGGCCATGAGCTGAGCT 916
DB 1145 TCTAGTTGGGCGGAGGTTTCTGAACCTGAGCAAGCTTACCTGGCCATGAGCTGAGCT 1204
QY 917 CGTTGTCGCGTCTTCACTGTTCTGAGTCTTCTTATGTTGCTGAGCTGAGTGC 976
DB 1205 CGTTGTCGCGTCTTCACTGTTCTTCTTATGTTGCTGAGCTGAGTGC----- 1257

QY 977 AGAGCAGCTCATCAACCCCTTTGGAGAGATGATGATTTTGAACCAACTGATTCGT 1036
DB 1258 ----- 1257
QY 1037 CGACGAAATTTGCAAGTGTCCCTGTTGGCTGTGATGATGACCAAGACCTGCTCG 1096
DB 1258 -----AGGTGTCCTGTTGGCTGTGATGATGATGACCAAGACCTGCTCG 1303
QY 1097 GATGAGCCGGAACATGATCTGGAATAAGCCCGGACCAAGCCCTTACACAGCTGCTTC 1156
DB 1304 GATGAGCCGGAACATGATCTGGAATAAGCCCGGACCAAGCCCTTACACAGCTGCTTC 1363
QY 1157 CGCCAGTTCCGTGAGCTCTTCTTATGAGCTCCACCTTCAACATCAGCTGAAACAAAG 1216
DB 1364 CGCCAGTTCCGTGAGCTCTTCTTATGAGCTCCACCTTCAACATCAGCTGAAACAAAG 1423
QY 1217 GGAATGGAAGTCCAGGCCCAATCAGAGAGAGAGAGAGATGTCACGCTGGCATCTTG 1276
DB 1424 GGAATGGAAGTCCAGGCCCAATCAGAGAGAGAGAGAGATGTCACGCTGGCATCTTG 1483
QY 1277 CGGTTCTAGGCTGACATGCCATGATCACAATCTCCAGGAGCAATCAAGAGCA 1336
DB 1484 CGGTTCTAGGCTGACATGCCATGATCACAATCTCCAGGAGCAATCAAGAGCA 1543
QY 1337 ACTACTGTGCCCAGAGAGGATCCCTTCTCAAGAGGCTGCCCCAAACCAAGAGC 1396
DB 1544 ACTACTGTGCCCAGAGAGGATCCCTTCTCAAGAGGCTGCCCCAAACCAAGAGC 1603
QY 1397 AGCCAAACAGAAAGTTAGGGGCCAGAGAGACAAAGAGCTGGAAGCTTAAAGCTGTGA 1456
DB 1604 AGCCAAACAGAAAGTTAGGGGCCAGAGAGACAAAGAGCTGGAAGCTTAAAGCTGTGA 1663
QY 1457 CGCTTCAAGTGTGCCCAGCTGATCAGAGGCAAGGCTTACAGTGCCTCCACAGAGCC 1516
DB 1664 CGCTTCAAGTGTGCCCAGCTGATCAGAGGCAAGGCTTACAGTGCCTCCACAGAGCC 1723
QY 1517 CCTCAGCCCACTCCCATGTTCTTCCCTTGAACCAATCAGGCGCTCAAGCTTCAAG 1576
DB 1724 CCTCAGCCCACTCCCATGTTCTTCCCTTGAACCAATCAGGCGCTCAAGCTTCAAG 1783
QY 1577 TGTCAAGGCAATGACACCAAGCAAAAGCTTAAAGCTGATGTTTGGGGCCAAAG 1636
DB 1784 TGTCAAGGCAATGACACCAAGCAAAAGCTTAAAGCTGATGTTTGGGGCCAAAG 1843
QY 1637 AAGTTTGAATGCTCTGAGAGAGAGAGAGAGGCTTGAATGAGACCCAGAAATATCTCA 1696
DB 1844 AAGTTTGAATGCTCTGAGAGAGAGAGAGAGGCTTGAATGAGACCCAGAAATATCTCA 1903
QY 1697 AGTGAAGAGAAACTGTGAGTTTAACTGACGGAATATGACAGAGATCCCGGAAATCA 1756
DB 1904 AGTGAAGAGAAACTGTGAGTTTAACTGACGGAATATGACAGAGATCCCGGAAATCA 1963
QY 1757 CCTGAAAGAACTTTGGAACATCAACCAACATACACATCACTCAAAAGATCAAT 1816
DB 1964 CCTGAAAGAACTTTGGAACATCAACCAACATCAACATCACTCAAAAGATCAAT 2023
QY 1817 GGAATCTTATGAGGCTTGAAGAAACAGGAGTGAAGCAATCTCAACCTGCTTCAATG 1876
DB 2024 GGAATCTTATGAGGCTTGAAGAAACAGGAGTGAAGCAATCTCAACCTGCTTCAATG 2083
QY 1877 GGAATCTTGAAGCAAGCAAGCTCTGACCTGTGTATACACAGAGAGACATGATCCAGTC 1936
DB 2084 GGAATCTTGAAGCAAGCAAGCTCTGACCTGTGTATACACAGAGAGACATGATCCAGTC 2143
QY 1937 ACAGCATAACAGCTGTCCACATGGAAGAGAGTGTCTTACCAACAGCTGATCAATGTTT 1996
DB 2144 ACAGCATAACAGCTGTCTTCACTGAAGAGATGTCTTACCAACAGCTGATCAATGTTT 2203
QY 1997 AGCTTAATAGATAAAATCCAGACTACTTCAAGCTTATGAGCTTATATCAATAAAG 2056
DB 2204 AGCTTAATAGATAAAATCCAGACTACTTCAAGCTTATATGCTTATATCAATAAAG 2263
QY 2057 TGTGAAAGCTAGACTAGACCATTTGAGAAACATTTAACTCAAGCTTGAATTCAGAGTGGG 2116

	78.8%;	Score 1756.4;	DB 9;	Length 1758;	
	Beat Local Similarity	99.9%;	Pred. No. 0;		
	Matches 1757;	Conservative	0;	Mismatches 1;	Indels 0; Gaps 0;

105 ATGACCATCACTTACAACAAGCCAAATGGCTTAATGCCCGCTTAGGCTTCCTTGCCCGCCTG 164

Db	1	ATGACCACTTACACAAAGCCAAAGTGGCTAATGAGCCGCTTAAGGCTCTTCCCGGCTG	60
QY	165	CTGCTGTGCTGGCGGGGAGACATCTACAGCTGCTATATGCGAGATTCTTAATCTTCTG	224
Db	61	CTGCTGTGCTGGCGGGGAGCATCTACAAAGCTGCTATATGCGAGATTCTTAATCTTCTG	120
QY	225	CTGCTGTCTATCATCATCCGCTTTATTTATATAGGCTGGGCTTCAGAGGAACAAACAGCTG	284
Db	121	CTGCTCATCTACATCATCCGCTTTATTTATATAGGCTGGGCTTCAGAGGAACAAACAGCTG	180
QY	285	ATGTTTGGAAATGCACTCTGTATTTGGCAGACAGTACATCCAGCTCATCCCAATTTCTTTC	344
Db	181	ATGTTTGGAAATGCACTCTGTATTTGGCAGACAGTACATCCAGCTCATCCCAATTTCTTTC	240
QY	345	GTGCTGGGCTTCTACGTGACGCTGGTCTGTACCCGCTGTGGAACAGTACGAAACCTG	404
Db	241	GTGCTGGGCTTCTACGTGACGCTGGTCTGTACCCGCTGTGGAACAGTACGAAACCTG	300
QY	405	CCGTGAGCCCGAACCCGCTCATAGAGCTGGTGTGCGGGCTTGGTGAAGGCAAGGACGAGCA	464
Db	301	CCGTGAGCCCGAACCCGCTCATAGAGCTGGTGTGCGGGCTTGGTGAAGGCAAGGACGAGCA	360
QY	465	GAGCGGCTGTGCGGGGCGCAGCTCATCCGTACAGCAACTGTGGCAAGTCTCATCTCTG	524
Db	361	GAGCGGCTGTGCGGGGCGCAGCTCATCCGTACAGCAACTGTGGCAAGTCTCATCTCTG	420
QY	525	CGACAGCTGACACCGCAGTCTACAGCGCTTCCCGAGCGCCAGCACTTGGTGAACA	584
Db	421	CGACAGCTGACACCGCAGTCTACAGCGCTTCCCGAGCGCCAGCACTTGGTGAACA	480
QY	585	GCGTTTATGATCTCGGCGAGAACACAGAGATTGGAGAACTGAGCTTACACACAACATG	644
Db	481	GCGTTTATGATCTCGGCGAGAACACAGAGATTGGAGAACTGAGCTTACACACAACATG	540
QY	645	TTCTGAGGAGCCCTGGGTGTGGTTTGGCCAAACGTCATGAAGAAGCGTGGCTTGAAGTGTGA	704
Db	541	TTCTGAGGAGCCCTGGGTGTGGTTTGGCCAAACGTCATGAAGAAGCGTGGCTTGAAGTGTGA	600
QY	705	ATCCGGGACCCCTATCTGCTTCAGAGCCCTGTGTAACGAGATGAACACTTGGGTACTGAG	764
Db	601	ATCCGGGACCCCTATCTGCTTCAGAGCCCTGTGTAACGAGATGAACACTTGGGTACTGAG	660
QY	765	TGTGGAACACTGTATGCTTACGACTGTGATATCCACCTGTGTATACACAGGTGTGTG	824
Db	661	TGTGGAACACTGTATGCTTACGACTGTGATATCCACCTGTGTATACACAGGTGTGTG	720
QY	825	ACTGTGGGGGTATCAGGCTTCTTCTGTGCTTGTCAAGTTGGAGCGGCACTTCTGAACCA	884
Db	721	ACTGTGGGGGTATCAGGCTTCTTCTGTGCTTGTCAAGTTGGAGCGGCACTTCTGAACCA	780
QY	885	GCCAAAGGCTTACCCCTGGGCATGAGCTGAGACCTGTGTGGCCGTTCTTACAGTTCTCTGAG	944
Db	781	GCCAAAGGCTTACCCCTGGGCATGAGCTGAGACCTGTGTGGCCGTTCTTACAGTTCTCTGAG	840
QY	945	TTCTTCTTCTATGTTGGCTGGCTGAAGGTGGAGAGCAAGCTCATTAACCCCTTTTGAAG	1004
Db	841	TTCTTCTTCTATGTTGGCTGGCTGAAGGTGGAGAGCAAGCTCATTAACCCCTTTTGAAG	900
QY	1005	GATGATGATGATTTTGGAGACCACTGATTTGTCGACAGAAATTTTGCAGAGTGTCCCTGTG	1064
Db	901	GATGATGATGATTTTGGAGACCACTGATTTGTCGACAGAAATTTTGCAGAGTGTCCCTGTG	960
QY	1065	GCTGTGATGTAGATGACACAGAGACTGTGCTCGATGTGAGCCGCAATGTACTGTGAATAG	1124
Db	961	GCTGTGATGTAGATGACACAGAGACTGTGCTCGATGTGAGCCGCAATGTACTGTGAATAG	1024
QY	1125	CCCGAGCCACAGCCCCCTTACACAGCTGTTCGCGCCAGTTTCGTGAGCTTCTTTATG	1184
Db	1021	CCCGAGCCACAGCCCCCTTACACAGCTGTTCGCGCCAGTTTCGTGAGCTTCTTTATG	1080
QY	1185	GCGTTCACCTTCAACATCAGGCTGAACAAAGAGAGATGAGATTCAAGCCCAATCAGAG	1244
Db	1081	GCGTTCACCTTCAACATCAGGCTGAACAAAGAGAGATGAGATTCAAGCCCAATCAGAG	1144

QY 1245 GACGAGGAGATGCTCAACGCTGATCATTTGGCCGCTTCTTAAAGCCCTGACGTCCCATGAT 1304
DB 1141 GAGGAGGAGATGCTCAACGCTGATCATTTGGCCGCTTCTTAAAGCCCTGACGTCCCATGAT 1200
QY 1305 GACCATCTCTCCAGGAGCAATCTCAAGAGCAAACTACTGTGGCCCAAGAGGAATCCCTT 1364
DB 1201 GACCATCTCTCCAGGAGCAATCTCAAGAGCAAACTACTGTGGCCCAAGAGGAATCCCTT 1260
QY 1365 CTCACAGAGGCTGCGCCAAAACCAAGAGCAAGCCCAAGAAAGTTAGGGGCAAGAA 1424
DB 1261 CTCACAGAGGCTGCGCCAAAACCAAGAGCAAGCCCAAGAAAGTTAGGGGCAAGAA 1320
QY 1425 GACAAAGAGGCTGGAAGCTTAAAGCTGTGAGCGCTTCAAGTCTGAGCCCACTATACG 1484
DB 1321 GACAAAGAGGCTGGAAGCTTAAAGCTGTGAGCGCTTCAAGTCTGAGCCCACTATACG 1380
QY 1485 AGGCAAGGCTTAAAGTGGCCCAAGAGCGCCCTCAAGCCCACTCCCATGTTCTTCCG 1544
DB 1381 AGGCAAGGCTTAAAGTGGCCCAAGAGCGCCCTCAAGCCCACTCCCATGTTCTTCCG 1440
QY 1545 CTAGAACCATCAGCGCCCTCAAGGCTTCAAGTGTCAAGGCAAGCAACCAAGAA 1604
DB 1441 CTAGAACCATCAGCGCCCTCAAGGCTTCAAGTGTCAAGGCAAGCAACCAAGAA 1500
QY 1605 AGCTTAAAGACTGTGAGTCTTGGGCGCAAGAAAGTTTGAATTTGCTCAGAGAGCGAT 1664
DB 1501 AGCTTAAAGACTGTGAGTCTTGGGCGCAAGAAAGTTTGAATTTGCTCAGAGAGCGAT 1560
QY 1665 GGGGCTTGTGATGAGAGCAAGGATCTCAAGTGTGAGAGGAGAACTGTGAGGTTTAA 1724
DB 1561 GGGGCTTGTGATGAGAGCAAGGATCTCAAGTGTGAGAGGAGAACTGTGAGGTTTAA 1620
QY 1725 CTGACGATATGACGAGGATCCCGAAATCCTCAAGAACTTTGAGCAATCAACA 1784
DB 1621 CTGACGATATGACGAGGATCCCGAAATCCTCAAGAACTTTGAGCAATCAACA 1680
QY 1785 ACCAATATACACTCACTCAAGATCATGAGATCTTATTTGGGCTTGGAAACAGG 1844
DB 1681 ACCAATATACACTCACTCAAGATCATGAGATCTTATTTGGGCTTGGAAACAGG 1740
QY 1845 GATGAGGACATTTCTTAA 1862
DB 1741 GATGAGGACATTTCTTAA 1758

RESULT 10
AX745964 1758 bp DNA linear PAT 14-MAY-2003
LOCUS AX745964
DEFINITION Sequence 1 from Patent WO03030922.
ACCESSION AX745964
VERSION AX745964.1 GI:30724619
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Steuernagel A., Bruenner G., Fritsch R., Eulenbergh K. and
Closek T.
TITLE Bestrophin and bestrophin homologous proteins involved in the
regulation of energy homeostasis
JOURNAL Patent: WO 03030922-A 1 17-APR-2003;
Develgen Aktiengesellschaft fuer Entwicklungsbiologische Forschung
(DE)
FEATURES
Source Location/Qualifiers
1..1758
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 78.7%; Score 1754.8; DB 6; Length 1758;

Best Local Similarity 99.9%; Pred. No. 0;
Matches 1756; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 105 ATGACCATCATTTACAAAGCAAGGCTATATGCGGCTTAAAGCTCTTCCGCTTG 164
DB 1 ATGACCATCATTTACAAAGCAAGGCTATATGCGGCTTAAAGCTCTTCCGCTTG 60
QY 165 CTGCTGTGCTGGCGGGGAGCATCTCAAGGCTGTATATGCGAGTTCTTAATCTTCTG 224
DB 61 CTGCTGTGCTGGCGGGGAGCATCTCAAGGCTGTATATGCGAGTTCTTAATCTTCTG 120
QY 225 CTGCTACTATCATATCCGCTTATTTATAGGTGCGCTTCAAGGAAAGCAAGCTG 284
DB 121 CTGCTACTATCATATCCGCTTATTTATATATAGGTGCGCTTCAAGGAAAGCAAGCTG 180
QY 285 ATGTTTGAAGAACTGACTGTGTATTTGAGAGCTACATCCAGCTCATATCCATTTCTTG 344
DB 181 ATGTTTGAAGAACTGACTGTGTATTTGAGAGCTACATCCAGCTCATATCCATTTCTTG 240
QY 345 GTGCTGGGCTTCTAAGTGTGAGAGCTGTGTGACCCGCTGTGTGAAACGATACGAAACCTG 404
DB 241 GTGCTGGGCTTCTAAGTGTGAGAGCTGTGTGACCCGCTGTGTGAAACGATACGAAACCTG 300
QY 405 CCGTGGCCGCAAGGCTCATGAGCTGTGTGCGGCTTGTGTGAAAGGCAAGGACGAGCA 464
DB 301 CCGTGGCCGCAAGGCTCATGAGCTGTGTGCGGCTTGTGTGAAAGGCAAGGACGAGCA 360
QY 465 GGCCTGGCTGTGAGGAGCAAGCTCATGAGCTGTGTGAGCAAGCTGTGTGAGCTATCTCTG 524
DB 361 GGCCTGGCTGTGAGGAGCAAGCTCATGAGCTGTGTGAGCAAGCTGTGTGAGCTATCTCTG 420
QY 525 CGAGGCTGAGCAAGGCTGTGAAAGCTTCCGAGGCGGCAAGCAAGCTGTGTGAGCA 584
DB 421 CGAGGCTGAGCAAGGCTGTGAAAGCTTCCGAGGCGGCAAGCAAGCTGTGTGAGCA 480
QY 585 GAGTTATGATCTCGGCAAGCAAGGCTGTGAAAGCTGTGAGCAAGCTGTGTGAGCA 644
DB 481 GAGTTATGATCTCGGCAAGCAAGGCTGTGAAAGCTGTGAGCAAGCTGTGTGAGCA 540
QY 645 TTCTGGGTGCTGTGAGGCTGTGAAAGCTGTGAAAGCTGTGAGCAAGCTGTGTGAGCA 704
DB 541 TTCTGGGTGCTGTGAGGCTGTGAAAGCTGTGAAAGCTGTGTGAGCAAGCTGTGTGAGCA 600
QY 705 ATCCGGAGCCCTATCTGTGCTGAGAGCTGTGAAAGCTGTGTGAGCAAGCTGTGTGAGCA 764
DB 601 ATCCGGAGCCCTATCTGTGCTGAGAGCTGTGAAAGCTGTGTGAGCAAGCTGTGTGAGCA 660
QY 765 TTGGAACCTGTATGCTGTGAGAGCTGTGAAAGCTGTGTGAGCAAGCTGTGTGAGCA 824
DB 661 TTGGAACCTGTGTGAGAGCTGTGAAAGCTGTGTGAGCAAGCTGTGTGAGCAAGCTGTGTGAGCA 720
QY 825 ACTGTGGGCTGTGAGAGCTGTGTGAGAGCTGTGTGAGAGCTGTGTGAGCAAGCTGTGTGAGCA 884
DB 721 ACTGTGGGCTGTGAGAGCTGTGTGAGAGCTGTGTGAGAGCTGTGTGAGCAAGCTGTGTGAGCA 780
QY 885 GCCAAGGCTTACCTGTGAGAGCTGTGTGAGAGCTGTGTGAGAGCTGTGTGAGCAAGCTGTGTGAGCA 944
DB 781 GCCAAGGCTTACCTGTGAGAGCTGTGTGAGAGCTGTGTGAGAGCTGTGTGAGCAAGCTGTGTGAGCA 840
QY 945 TTCTTCTTATATTTGCTGTGAGAGCTGTGTGAGAGCTGTGTGAGAGCTGTGTGAGCAAGCTGTGTGAGCA 1004
DB 841 TTCTTCTTATATTTGCTGTGAGAGCTGTGTGAGAGCTGTGTGAGAGCTGTGTGAGCAAGCTGTGTGAGCA 900
QY 1005 GATGATGATATTTTGAAGCAAGCTGTGTGAGAGCTGTGTGAGAGCTGTGTGAGCAAGCTGTGTGAGCA 1064
DB 901 GATGATGATATTTTGAAGCAAGCTGTGTGAGAGCTGTGTGAGAGCTGTGTGAGCAAGCTGTGTGAGCA 960
QY 1065 GCTGTGATGATGATGAGCAAGCTGTGTGAGAGCTGTGTGAGAGCTGTGTGAGCAAGCTGTGTGAGCA 1124
DB 961 GCTGTGATGATGATGAGCAAGCTGTGTGAGAGCTGTGTGAGAGCTGTGTGAGCAAGCTGTGTGAGCA 1020
QY 1125 CCGAGCCAGAGCCCTTACAGAGCTGTGTGAGAGCTGTGTGAGAGCTGTGTGAGCAAGCTGTGTGAGCA 1184

Db	1021	CCCGAGCCACAGCCCCCTTACAGAGCTGCTTCGCCGCCAGTTCCGTGAGCCTCTTATATG	1080
Qy	1185	GGCTCCACCTTCAACATCAGCCTGAAACAAAGAGAGATGAGTTCAGGCCCATCAGAG	1244
Db	1081	GGCTCCACCTTCAACATCAGCCTGAAACAAAGAGAGATGAGTTCAGGCCCATCAGAG	1140
Qy	1245	GAGCAGAGAGATGCTCAACGCTGGCATCTTGGCCGCTTCCTAGGCTGACAGTCCATGAT	1304
Db	1141	GAGCAGAGAGATGCTCAACGCTGGCATCTTGGCCGCTTCCTAGGCTGACAGTCCATGAT	1200
Qy	1305	CACCATCTCTCCAGGCGCAATCTCAAGACCAACTACTGTGTGGCCCAAGAGGGAATCCCTT	1364
Db	1201	CACCATCTCTCCAGGCGCAATCTCAAGACCAACTACTGTGTGGCCCAAGAGGGAATCCCTT	1260
Qy	1365	CTCCAGCAGGCGCTGCCCAAAACCAACAGGAGCAGCAAGACGATTAAGGAGCAGAGAA	1424
Db	1261	CTCCAGCAGGCGCTGCCCAAAACCAACAGGAGCAGCAAGACGATTAAGGAGCAGAGAA	1320
Qy	1425	GACAAACAGGCGCTGGAAGCTTAAAGCTGTGTGACGCTTCAAGTCTGCCCACATGATCAG	1484
Db	1321	GACAAACAGGCGCTGGAAGCTTAAAGCTGTGTGACGCTTCAAGTCTGCCCACATGATCAG	1380
Qy	1485	AGGCCAGGCTACTACAGTGTCCCAACAGAGCGCCCTCAGGCCCATCCCATTTCTTCCC	1544
Db	1381	AGGCCAGGCTACTACAGTGTCCCAACAGAGCGCCCTCAGGCCCATCCCATTTCTTCCC	1440
Qy	1545	CTAAGACCATCAGGCGCTGAAAGCTTCAAGCTGTGTGACGAGCATGACACCAAGACAA	1604
Db	1441	CTAAGACCATCAGGCGCTGAAAGCTTCAAGCTGTGTGACGAGCATGACACCAAGACAA	1500
Qy	1605	AGCTTAAAGACTGTGAGTCTTGAGGCGCCAAAGAAATTTGTAATGCTCTCAGAGAGCGAT	1664
Db	1501	AGCTTAAAGACTGTGAGTCTTGAGGCGCCAAAGAAATTTGTAATGCTCTCAGAGAGCGAT	1560
Qy	1665	GGGGCCCTTGATGAGCAGCCCAAGATCTTAAGTGAAGAGAGAAACTGTGAGTTTAC	1724
Db	1561	GGGGCCCTTGATGAGCAGCCCAAGATCTTAAGTGAAGAGAGAAACTGTGAGTTTAC	1620
Qy	1725	CTGACGATATGCGAGAGATCCCGGAAATACCTCAAGAAACCTTTGGAAACATTCACCA	1784
Db	1621	CTGACGATATGCGAGAGATCCCGGAAATACCTCAAGAAACCTTTGGAAACATTCACCA	1680
Qy	1785	ACCAACATACACACTACACTCAAGATCAACATGATCTTATTTGGGCTTTGAAACAGG	1844
Db	1681	ACCAACATACACACTACACTCAAGATCAACATGATCTTATTTGGGCTTTGAAACAGG	1740
Qy	1845	GATGAGCAGCATTCCTTAA	1862
Db	1741	GATGAGCAGCATTCCTTAA	1758
RESULT 11			
BC041664		2435 bp mRNA linear PRI 07-OCT-2003	
LOCUS		Homo sapiens vitelliform macular dystrophy (Best disease,	
DEFINITION		bestrophin), mRNA (cDNA clone MGC:47884 IMAGE:5194649), complete	
cds			
ACCESSION	BC041664		
VERSION	BC041664.1	GI:27371319	
KEYWORDS	MGC.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE			
AUTHORS			

TITLE	ABRAMSON, R.D., MULLAHY, S.J., BOSAK, S.A., MCEWAN, P.J., MCKERNAN, K.J., MALEK, J.A., GUNARATNE, P.H., RICHARDS, S., MCKLEY, K.C., HALE, S., GARCIA, A.M., GAY, L.J., HULYK, S.W., VILLALON, D.K., MUZYNY, D.M., SODERGREN, E.J., LU, X., GIBBS, R.A., RAHEY, J., HELTON, E., KETEMAN, M., MADAN, A., RODRIGUES, S., SANCHEZ, A., WHITING, M., MADAN, A., YOUNG, A.C., SHEVCHENKO, Y., BOUFFARD, G.G., BLAKESLEY, R.W., TOUCHMAN, J.W., GREEN, E.D., DIKEON, M.C., RODRIGUEZ, A.C., GRIMWOOD, J., SCHMUTZ, J., MYERS, R.M., BUTTERFIELD, Y.S., KRZYWINKI, M.I., SKALSKA, U., SMALUS, D.E., SCHNERCH, A., SCHEIN, J.E., JONES, S.J. and MARR, M.A.
JOURNAL	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
MEDLINE	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED	22388257
REFERENCE	2 (bases 1 to 2435)
AUTHORS	Straubberg, R.
TITLE	Direct Submission
JOURNAL	Submitted (20-DEC-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT	Contact: MGC help desk Email: cgabs-remail.nih.gov Tissue Procurement: Life Technologies, Inc. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL) DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland, Web site: http://www.nisc.nih.gov/ Contact: nisc.mgc@nih.gov Ahter, N., Ayala, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaspi, R., Madu, Q.L., Maiello, C., Maekari, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantirpop, S., Thomas, P.J., Touchman, J.W., Turgeson, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: http://image.llnl.gov Series: IRAC Plate: 82 Row: d Column: 4. Location/Qualifiers 1. 2435 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="MGC:47884 IMAGE:5194649" /tissue_type="Brain, adult, 6 pooled whole brains" /clone_lib="NIH_MGC_114" /lab_host="DH10B" /note="Vector: pCMV-SPORT6" 1. 2435 /gene="VMD2" /note="synonyms: BMD, BEST" /db_xref="LocusID:7439" /db_xref="MIM:153700" 87..1901 /gene="VMD2" /codon_start=1 /product="VMD2 protein" /protein_id="AA41664.1" /db_xref="GI:27371320" /db_xref="LocusID:7439" /db_xref="MIM:153700" /translation="MEKTLVCDSPYLVYVAVYSPVFGVYTLVTRKMYNXYLPPMDRLMSLVSGPREGDEBGRILRLTILRYALGNLILRSVTAAYKRPSPQHYVQAFMTAEKHKOLEKSLPINMEFVPPWVFWANLSMKMLGRIIDPILLOSLLNEKMLRTQCGHLYAVDWISIPLVYTVVTAIVYFELICLVGRQFLNPAKYVPGHIEDLVVVPFTPLGPFYVGMKVAEQLINPGEDEDDFETWIVDRNLQVSLAVDEKHQDLPRMEPDM
CDs	

YMKRPEQPYTAASQFRVSEFWGTFENISLNKEEMEPQNEDEDAHAGTIGRPL
GLSDHHPHPRANSRTKLMPKRESLHBLPRNHKRAKONUNGQDNKANKLKYANDA
FKSAPLYORPOYVSAPOPLSPMPPEPLPSAPSKLHSTGIDTIDKSLKVSQAK
KSPFLSESGALMEHPEVSOVRKYVERLUTMPPELPEHMLKEPLKESQSPVNIHTLK
DHMDPYALNRSYSLHNOGCTALCPELHSLSLPFLHFLNFHHCOSTLDRPAL
AMGILYLTFTGILGKSGPFLTSPWYHPEDFLQPSGR" 90..1022
misc_feature
/gene="VMD2"
/note="DUF289; Region: Putative membrane protein. This family called family 8 in. may be a transmembrane protein The specific function of this protein is unknown"
/db_xref="CD:pfam01062"

ORIGIN

Query Match 71.1%; Score 1584; DB 9; Length 2435;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1587; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 257 GCTGGCCCTCAGGAGAAACAAGCTGATGTTTGAAACTGACTGTGTATTTGGCAG 316
Db 59 GCTGGCCCTCAGGAGAAACAAGCTGATGTTTGAAACTGACTGTGTATTTGGCAG 118
QY 317 CTACATCAGCTCATCCCATTTCTTCGNGCTGGCTCTAAGTACGCTGTGTAC 376
Db 119 CTACATCAGCTCATCCCATTTCTTCGNGCTGGCTCTAAGTACGCTGTGTAC 178
QY 377 CCGCTGTGGAACCAAGTACGAGAACCTGCGTGGCCGACCGCTCATAGCCTGTGTC 436
Db 179 CCGCTGTGGAACCAAGTACGAGAACCTGCGTGGCCGACCGCTCATAGCCTGTGTC 238
QY 437 GGGCTTCGTGAGGCAAGGAGGAGGCGCTGCGGCGGCGCAAGCTCATAGCCTGTGTC 496
Db 239 GGGCTTCGTGAGGCAAGGAGGAGGCGCTGCGGCGGCGCAAGCTCATAGCCTGTGTC 298
QY 497 CCGCAACTGGGCAACGTGCTCATCTCGGACGCTGACCGCAGTCTCAAGCGCTT 556
Db 299 CCGCAACTGGGCAACGTGCTCATCTCGGACGCTGACCGCAGTCTCAAGCGCTT 358
QY 557 CCCCAGCCGCGAGCACTGGTGCAAGGAGGCTTTATGACTCCGCAACAACAAGCTT 616
Db 359 CCCCAGCCGCGAGCACTGGTGCAAGGAGGCTTTATGACTCCGCAACAACAAGCTT 418
QY 617 GAGAAACTGAGCCTTACCAACAATGTTCTGGGTGCGCTGAGTGTGGTGGCACT 676
Db 419 GAGAAACTGAGCCTTACCAACAATGTTCTGGGTGCGCTGAGTGTGGTGGCACT 478
QY 677 GTCAATGAGGCGTGGCTTGAAGCTCGAATCGGAGCCTTATCTGCTCGAAGCCTGCT 736
Db 479 GTCAATGAGGCGTGGCTTGAAGCTCGAATCGGAGCCTTATCTGCTCGAAGCCTGCT 538
QY 737 GAAAGAGTAAACCTTGTGCTACTCAGTGTGACCTGTATGAGCCCTAACAAGCTT 796
Db 539 GAAAGAGTAAACCTTGTGCTACTCAGTGTGACCTGTATGAGCCCTAACAAGCTT 598
QY 797 TATCCCACTGTGTATACAGAGTGTGACTGTGGCGGTGTAGAGCTTCTCTGACTTG 856
Db 599 TATCCCACTGTGTATACAGAGTGTGACTGTGGCGGTGTAGAGCTTCTCTGACTTG 658
QY 857 TCTAGTGGGCGGAGTGTCTGAACCAAGCAAGGCTTACCTGGCAATGAGTGAAGCTT 916
Db 659 TCTAGTGGGCGGAGTGTCTGAACCAAGCAAGGCTTACCTGGCAATGAGTGAAGCTT 718
QY 917 CGTTGTGCGGCTTCAAGTGTCTGAGTGTCTTCTAATGTTGGCTGTGAAGTGGC 976
Db 719 CGTTGTGCGGCTTCAAGTGTCTGAGTGTCTTCTAATGTTGGCTGTGAAGTGGC 778
QY 977 AGAGAGCTCATAAACCTTTTGAAGAGTATGATGATTTTGAAGCAACTGATGAT 1036
Db 779 AGAGAGCTCATAAACCTTTTGAAGAGTATGATGATTTTGAAGCAACTGATGAT 838
QY 1037 CGACAGAAATTTGACAGTGTCCCTGTTGGCTGTGAGTGAAGCAACGAGACCTGCTCG 1096
Db 839 CGACAGAAATTTGACAGTGTCCCTGTTGGCTGTGAGTGAAGCAACGAGACCTGCTCG 898

QY 1097 GATGAGCCGGAACATGTAAGTGAATTAAGCCGAGCCACAGCCCCCTACACAGTCTTC 1156
Db 899 GATGAGCCGGAACATGTAAGTGAATTAAGCCGAGCCACAGCCCCCTACACAGTCTTC 958
QY 1157 CGCCCACTTCCTGTAGAGCTTCCTTATGAGGCTTCACCTTCAAGATCAGCTGAAACAAGA 1216
Db 959 CGCCCACTTCCTGTAGAGCTTCCTTATGAGGCTTCACCTTCAAGATCAGCTGAAACAAGA 1018
QY 1217 GGAAGTGAAGTTCAGGCCCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1276
Db 1019 GGAAGTGAAGTTCAGGCCCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1078
QY 1277 CCGCTTCTTATGAGGCTGTGATCCATGATCAATCTCTCCAGGCAACTCAAGACCAA 1336
Db 1079 CCGCTTCTTATGAGGCTGTGATCCATGATCAATCTCTCCAGGCAACTCAAGACCAA 1138
QY 1337 ACTACTGTGCGCAAGAGGGAATCCCTTCTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1396
Db 1139 ACTACTGTGCGCAAGAGGGAATCCCTTCTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1198
QY 1397 AGCCAAACAGAACGTTAGGGGCGCAGAGAACAAAGGCTGAAAGCTTAAAGCTGTGGA 1456
Db 1199 AGCCAAACAGAACGTTAGGGGCGCAGAGAACAAAGGCTGAAAGCTTAAAGCTGTGGA 1258
QY 1457 GCGCTTCAAGTGTGAGCCCACTGTATGAGAGGCGAGGCTTCAAGTGTGAGGCGC 1516
Db 1259 GCGCTTCAAGTGTGAGCCCACTGTATGAGAGGCGAGGCTTCAAGTGTGAGGCGC 1318
QY 1517 CCTAGCCCCACTCCCATTTCTTCCCTTAGAACATCAAGGCGGTCAAAAGCTTCAAG 1576
Db 1319 CCTAGCCCCACTCCCATTTCTTCCCTTAGAACATCAAGGCGGTCAAAAGCTTCAAG 1378
QY 1577 TGTCAAGGAGTATGACCAAAAGCAAAAGCTTAAAGCTGATGTTGTGGGCGCAAGA 1636
Db 1379 TGTCAAGGAGTATGACCAAAAGCAAAAGCTTAAAGCTGATGTTGTGGGCGCAAGA 1438
QY 1637 AAGTTTGAATTTGCTCTCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1696
Db 1439 AAGTTTGAATTTGCTCTCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1498
QY 1697 AGTGAAGAGGAAACTGTGAGTGAATTAAGTGAAGTATGAGGAGTATGAGGAGTATGAG 1756
Db 1499 AGTGAAGAGGAAACTGTGAGTGAATTAAGTGAAGTATGAGGAGTATGAGGAGTATGAG 1558
QY 1757 CCTCAAAACCTTTGGAACATCAACCAACCAATACATCACTCAAAAGTCAAT 1816
Db 1559 CCTCAAAACCTTTGGAACATCAACCAACCAATACATCACTCAAAAGTCAAT 1618
QY 1817 GGATCCTTATTTGGGCTTTGGAAGGAGGAGT 1848
Db 1619 GGATCCTTATTTGGGCTTTGGAAGGAGGAGT 1650

RESULT 12
LOCUS BD132659 1263 bp DNA linear PAT 18-SEP-2002
DEFINITION Secreted proteins and polynucleotides encoding them.
ACCESSION BD132659
VERSION BD132659.1 GI:23227604
KEYWORDS JP 2002504822-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Jacobs K., McCoy J.M., Lavallie B.R., Racie L.A., Treac M.,
Spaulding V., Agostino M.J., Howes S.H. and Fechtel K.
TITLE Secreted proteins and polynucleotides encoding them
JOURNAL JOURNAL
GENETICS INSTITUTE INC
PN JP 2002504822-A/2
PD 12-FEB-2002

	Pf	08-JUN-1998 JP 1999S03038								US /09Z722 PT
	PR	11-JUN-1997 US 08/68231D_05-JUN-1998 US								KENNETH JACOBS,JOHN M MCCOY,EWARD R LAVALLE,LISA A RACIE, PI MAURICE TREACY,
	PI	VIKKI SAULJINDING,MICHAEL J AGOSTINO,STEVEN H HOWES,KIM FECHTEL								
	PC	C12N15/12,C07K14/47,A61K38/17								
	CC	Strandedness: Double;								
	CC	Topology: Linear;								
	FH	Key Location/Qualifiers.								
FEATURES	source	/organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606"								
ORIGIN										
Query Match		56.0%; Score 1247.2; DB 6; Length 1263;								
Best Local Similarity		99.8%; Pred. No.6.5e-27;								
Matches 1249; Conservative		0; Mismatches 3; Indels 0; Gaps 0;								
Oy	978	GAGCAGCTCATCAACCCTTTGAGAGGATGATGTATTGGACCAACTGAATTGC	1037							
Dn	1	GAGCAGCTCATCAACCCTTTGAGAGGATGATGTATTGGACCAACTGAATTGC	60							
Oy	1038	GACAAGAAATTTTGACAGTTCCTCGTGTGCTGTGATGTAAGAACCAAGACTGCTCGG	1097							
Dn	61	GACAAGAAATTTTGACAGTTCCTCGTGTGCTGTGATGTAAGAACCAAGACTGCTCGG	120							
Oy	1098	ATGAGAAGCGGACATGTATACTGGAAATPAGGCCGAGCACAGCCCCCTTACACAGCTGCTCC	1157							
Dn	121	ATGAGAAGCGGACATGTATACTGGAAATPAGGCCGAGCACAGCCCCCTTACACAGCTGCTCC	180							
Oy	1158	GCCCAGTTCGGTCGAGCTCTCTTATGGGCTCCACCTTCAAATCAGGCTGAAACAAGAG	1217							
Dn	181	GCCCAGTTCGGTCGAGCTCTCTTATGGGCTCCACCTTCAAATCAGGCTGAAACAAGAG	240							
Oy	1218	GAGATGAGATTCACAGCCCCAATCAGAGAGACGAGAGGATGCTCAAGCTGTGATCATTTGGC	1277							
Dn	241	GAGATGAGATTCACAGCCCCAATCAGAGAGACGAGAGGATGCTCAAGCTGTGATCATTTGGC	300							
Oy	1278	CGCTTCCTCAGGSCGTGCACTGCCATATACCATCCTCCGAGGGCAAACTCAAGAGCCA	1337							
Dn	301	CGCTTCCTCAGGSCGTGCACTGCCATATACCATCCTCCGAGGGCAAACTCAAGAGCCA	360							
Oy	1338	CTACTGTGGCCCAAGAGGAAATCCCTTCTCCACGAGGGCTTGCCCAAAAACAAGAGCA	1397							
Dn	361	CTACTGTGGCCCAAGAGGAAATCCCTTCTCCACGAGGGCTTGCCCAAAAACAAGAGCA	420							
Oy	1398	GCCAAACAAGACGTTAAGGGGCCAGAGAACAACAAGGCTCTGAAGCTTAAGCTGTGGAC	1457							
Dn	421	GCCAAACAAGACGTTAAGGGGCCAGAGAACAACAAGGCTCTGAAGCTTAAGCTGTGGAC	480							
Oy	1458	GCGTTCAAGTCGTGGGCCCATGTGATCAGAGGGCCAGGCTACTACAGTGGCCCAAGCGGCC	1517							
Dn	481	GCGTTCAAGTCGTGGGCCCATGTGATCAGAGGGCCAGGCTACTACAGTGGCCCAAGCGGCC	540							
Oy	1518	CTCAGCCCCCATCTCCATGTTCTTCCCTCCTAGAACCATCAGCGCGCTCAAGCTTCACAGT	1577							
Dn	541	CTCAGCCCCCATCTCCATGTTCTTCCCTCCTAGAACCATCAGCGCGCTCAAGCTTCACAGT	600							
Oy	1578	GTCACAGGCATAGACACCAAAACAACAAAGCTTTAAAAAGTGTGAGTTCTGGGGCCAAAGAA	1637							
Dn	601	GTCACAGGCATAGACACCAAAACAACAAAGCTTTAAAAAGTGTGAGTTCTGGGGCCAAAGAA	660							
Oy	1638	AGTTTTGAATTTGCTCTCAGAGAGGGAATGAGAGGACCCAGAAAGTATCTCAA	1697							
Dn	661	AGTTTTGAATTTGCTCTCAGAGAGGGAATGAGAGGACCCAGAAAGTATCTCAA	720							
Oy	1698	GTCAGAGAGAAAACGTGTGAGTTTAACTGACGGAATATGCCAGAGANTCCCGAAATATAC	1757							
Dn	721	GTCAGAGAGAAAACGTGTGAGTTTAACTGACGGAATATGCCAGAGANTCCCGAAATATAC	780							
Oy	1758	CTCAAAGAACCCTTGGAAACAATCACCAACCAATATACACTACACTCAAAGATCATATG	1817							

[illegible]

OY	1209	AACAAAGGAGAAATGAGAGTTCCAGGCCCAATCAGAGAGACAGAGAGAGATGCTCAGCTG--	1266
Db	1141	AGGAAGAGAACTTGAAGCTTTGGTCAAAAGAGGAGGCTGACACCGAATAAGAAAGAGAT	1200
OY	1267	-----GCATCATTTGGCCGCTTCTCTAGGCTCGCAGTCCCATGATCACCATCTGCC	1316
Db	1201	GCGTATTAGCAGCAGCCATGAGCTGTCTTTTAAGACTGCACCCAAAACCTACCATCTTCC	1266
OY	1317	AGGGCAAACTCAAGGACCAAACTACTGTGTGGCCCAAGAGGAATCCCTTCTCAGAGAGGC	1376
Db	1261	TTGAAGAAGCTTAAAGACCAAACTATTGTGTTTAAAG-----ATCCCTCTCTGAAAGGC	1314
OY	1377	CTGCCAAAAACACAGGAGCAGCCAAACAGAACTTTAAGGGCCAGGAAGACAAAGAGCC	1436
Db	1315	CAGGTGAAGGATGCCAACACAGAAAAACAGAAAG-----ATGTC	1353
OY	1437	TGGAAGCTTAAAGCGTGTGGACGCCCTTAAAGCTGGCCCACTGTATCAGAGGCCAGGCTAC	1496
Db	1354	TGGAATTTAAAGGCGTGTGACTCTTTGTAATATGTGTTCCAAAGCTTTTAAAGGAGAGGCTCC	1413
OY	1497	TACAGTCCCCACAGACGCCCTCTAAGCCCACTCC	1531
Db	1414	CATTGTGGCCACAGGACACCCAGACGCCACCTTAC	1448

[illegible]

REFERENCE	1 (bases 1 to 1916)
AUTHORS	Petrushkin, K., Caskey, T.C., Metzker, M. and Wadelius, C.
TITLE	Beet's macular dystrophy gene
JOURNAL	Patent: JP 2002504559-A 26 12-FEB-2002;
	MERCK & CO INC, CLABS WADELIIUS
COMMENT	OS Mus musculus (mouse)

```
PC C07K16/18, C07K14/47, C12N5/10, C12N15/09, C12P19/34, C12Q1/68// PC
C12P21/08,
PC C12N5/00, C12N15/00
CC Best's macular dystrophy gene
FH Key Location/Qualifiers
FT source 1..1916
FT /organism='Mus musculus (mouse)'.
FEATURES
source 1..1916
```

ORIGIN

Query Match	39.1%;	Score 871.2;	DB 6;	Length 1916;
Best Local Similarity	77.1%;	Pred. No. 6.8e-186;		
Matches 116; Conservative	0;	Mismatches 293;	Indels 39;	Gaps 3;

QY 96 TGGCCGAGCATGACCATCACTTAACAAGCAAGTGGCTTAATGCCGCTTAGCGTCCTTC 155

Db 2 TGGCAAGCAGACTATCACTATCACTTAACAAGTAAGTACCAATGCCGCGCTCGATTCGCTTC 61

QY 156 TCCCGCGCTGCTGCTGTGTGGCGGGAGCATTTACAAGCTGTATATGACGAGTTCTTA 215

Db 62 TGGTCCCTCCCTCGCTGTCTGGCGAGGAGCATTTACAAGTGGTGTATGAAATATCTT 121

QY	216	ATTTCTCTGCTGCTGCTACTCATCATCCGCTTTATTTATTAAGCTGGCCCTCACGGAAGAA	275
Db	122	GTCTTATATATCTCTACTATTTCATCCCGTGACTCTACAGATGGTTCTCTCGATGTAT	181
QY	276	CACAGCTCATGTTTATAGAAACTGATCTGTATTTGAGCAGGCTACATCCAGCTCATCCCC	335
Db	182	CAGCACTCTTGTTTGAGAAAGCTGGCTCTGTACTGAGCAGCTACTACTTACGCTCATCCCT	241
QY	336	ATTTCCTTGTGTGGGCTTCTTACGTACGCTGGTCTGACCCGCTGGTGGAAACCACTAC	395
Db	242	ATATCTTGTGTTCTGGGTTTCTATGTTTATCATTTGGTGGTGAAGCCGCTGGTAGCCAGTAC	301
QY	396	GAGAACCTCCGCGGCGCCAGCCGCTCATGAGCCGTGGTGGGGGCTTGGTCCAGAGGCAAG	455
Db	302	GAGAACTTGCCTGGCCGACCCGCTCATGATTCAGAGTCTTACGCTTCTGAGAGGGCAAG	361
QY	456	GACGAGCAAGGCGGAGCTGTGCGAGCGCACGCTCATCCGTACGCGAACCTGGGCAACGTG	515
Db	362	GATGAGGAAGGCGGTTTGTGTGGCGGACGCTCATCCGCTACGCCATCCTGGGCCAAGTG	421
QY	516	CTCATCTCGGAGAGGCTAGACACCGCAGTCTTACAGCGCTTCCCGAGCGCCAGACCTG	575
Db	422	CTCATCTCGGAGAGATCAGACCTCGGCTTACAGCGCTTCCCACTTTCACCACTCG	481
QY	576	GTGCAAGCAGGCTTATGACTCCGSCAAGAACACAGCAGTTGGAGAAACTGAGCCTACCA	635
Db	482	GTGCTAGCAGGTTTATATGACCATGGGAACTATAGAGTTGCAAGATGGGCTTACCA	541
QY	636	CACACATGTTCTGGGTGTCCTGGGTGTGTTTGCCAACTGTCAATGAAAGCGTGCTT	695
Db	542	CACACATATCTGGGTGTCCTGGGTGTGTTTGCCAACTGTCAATGAAAGCGCTTACTT	601
QY	696	GGAGGTCAATCCGGGAGCCCTATCCCTGTGTCAGAGCCGCTGGAAGAGATGAACACTTG	755
Db	602	GGAGGTCAATCCGGGAGACCGTCTGTCTCCAGAGCTTGATGAAAGAGGTGTATCTTG	661
QY	756	CGTACTCAGTGTGACACCTGTATGCCCTACGACTGATTAAGTATCCACTGGTGTATACA	815
Db	662	CGTACTCAGTGTGACAGCTGTATGCCCTACGACTGATTAAGTATCCACTGGTGTATACA	721
QY	816	CAGGTGTGACTGTGGCGGTGTAACGTTCTTCTGACTTGTGTCTAAGTTGGGCGGCACTT	875
Db	722	CAGGTGTGACTGTGGCGATATTAAGCTTTTCTTGTGATCTTGAATCGGAGGCAAGTTT	781
QY	876	CTGAACCCAGCAAGGCTTACCCTGGGCACTAGCTGGACTCGTGTGCCGCTCTTACG	935
Db	782	CTGAACCCAAACAAAGACTACCCAGGCAATGAAATGATCTGGTTGTGCCCTGTCTTACA	841
QY	936	TTCTCTGCAATTCTTCTTATGTGTGGCTGAGGTGGAGAGCAAGCTCATCAACCC	995
Db	842	ATCTGTGCAATCTTATTTCTATCATGTGGCTGGCTGAAGGTGGGCAAGCAAGCTCATCAACCC	901
QY	996	TTTGAAGAGATGATGATTTTGAACCAACCTGATTTGTGCAAGAAATTTGCAAGTG	1055
Db	902	TTTGGGAGAGATGATGATTTTGAACCTTATCTGATCATTTGACAGAAACCTTCAAGTG	961
QY	1056	TTCTGTGTGCTGTGATGATGATGACACAGAGACTGTGCTGGATGAGGCCGGAATGTAC	1115
Db	962	TTCTGTGTGTCCGGATGAGATGACACAGAACTTGGCTCCATGGAACGTGACATGTAC	1021
QY	1116	TGGAATTAAGCCCAAGCCACAGCCCCCTTACACAGCTGTTCGCGCCAGTTCCGTGAGCC	1175
Db	1022	TGGAACAGGCAAGGCTTCAAGCCGCTTACACAGCTGTTCGTCCAGGTCTCGCGGCAT	1081
QY	1176	TCTTTATGGGCTCACTTCAACATCAGCCTTAAACAAAGAGATGAGTTCCAGCCC	1235
Db	1082	TCTTTATGGGCTTCACTTCAACATCAGCCTTAAAGAAAGAAACTTAAAGCTTTGGTCA	1141
QY	1236	AATTCAGAGACGAGAGAGATGCTCAAGCTG-----GCATCATTTGGCCGCTT	1283
Db	1142	AAAGAGAGGCTTACACGATTAAGAAAGAGAGTGCTATAGCAGACCAATAGGCTGTCTT	1201

```

QY 1284 CTAGGCTGCGATGCCATGATCAACATCTTCCAGGGCAAACTCAAGACCAACTACTG 1343
Db 1202 TTAGGACTGCACCAAAAATACTATCTTCCCTTGAAAGACTTAAAGACCAAACTATTG 1261
QY 1344 TGGCCCAAGAGGGAATCCCTTCTTCACAGAGGCTGCCCCAATAACCAAGAGCAGCCAA 1403
Db 1262 TGTTCATAAG-----ACCCCTCTCTCGAAGGCCAGTAAAGATGCCAAACAGAAAAAC 1315
QY 1404 CAGAACGTTAGGGGCCAGAAAGACAAGGCTGAAAGCTTAAAGGCTGTGAGCGCCTTC 1463
Db 1316 CAGAAAG-----ATGTCGGAATTAAAGGCTGTGACTTCTTG 1354
QY 1464 AAGTCTGCCCCACATGATACAGAGGCCAGCTACTACAGTCCCCACAGACCCCTCAGC 1523
Db 1355 AATGTGTTCCAAAGTTTAAGAGAGAGGCTCCCATTTGTGCCCCACAGGCAACCCAGCAGC 1414
QY 1524 CCCACTCC 1531
Db 1415 CACCTTAC 1422

```

Search completed: March 26, 2005, 07:38:20
 Job time : 9510.76 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2005, 22:43:02 ; Search time 7168.13 Seconds
(without alignments)
11836.456 Million cell updates/sec

Title: US-09-622-964A-2
Perfect score: 2229
Sequence: 1 caggaggtccaccagccta.....aaaaaaaaaaaaaaaa 2229

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479086

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : EST.*
1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_est4.*
5: gb_est5.*
6: gb_est6.*
7: gb_est7.*
8: gb_est8.*
9: gb_est9.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	777.6	34.9	963	5	B0879880 AGENCOURT
2	756.2	34.4	881	5	CD518675 AGENCOURT
3	754.2	33.8	1699	3	AK006549 Mus muscu
4	681.8	30.6	711	5	BK095540 BX095540
5	657	29.5	676	5	BT41926 UI-E-E01-
6	651.6	29.2	659	5	BT730894 UI-E-E01-
7	648.8	29.1	674	5	BT731809 UI-E-E01-
8	641.8	28.8	756	7	CO396179 AGENCOURT
9	640.4	28.7	666	4	BT480798 H2RPE-043
10	638	28.6	639	6	CA389968 cs104h03.
11	631.2	28.3	651	6	CA397981 cs98c05.Y
12	626.6	28.1	656	5	BT731149 UI-E-E01-
13	582.4	26.1	792	4	BM663028 UI-E-E01-
14	577	25.9	772	4	BM610951 60130362
15	575	25.8	585	4	BM707948 UI-E-E01-
16	565.2	25.4	593	2	BE385296 601277572
17	528.8	23.7	735	4	BT756228 BM691456 UI-E-E01-
18	525.8	23.6	537	4	BM691456 UI-E-E01-
19	522	23.4	522	4	BM707649 UI-E-E01-
20	508	22.8	526	4	BM685501 UI-E-E01-
21	495.4	22.2	497	6	CA391418 cs13a10.Y
22	492.4	22.1	930	5	BQ436824 AGENCOURT
23	492.2	22.1	545	6	CA395098 cs60c05.Y
24	491	22.0	491	7	CV024947 2492 Full

25	485.2	21.8	566	4	BM718338 UI-E-E01-
26	476.4	21.4	582	5	BP379588 BP379588
27	475.4	21.3	503	4	BM685122 UI-E-E01-
28	475.4	21.3	508	5	BM932117 UI-E-E01-
29	470.8	21.1	474	1	AI190190 qd37c03.X
30	469.6	21.1	492	5	BT726009 UI-E-E01-
31	467	21.0	484	5	BT741783 UI-E-E01-
32	464.8	20.9	469	1	AA307119 EST178031
33	449.2	20.2	455	4	BG013943 IL5-GN024
34	444.6	19.9	601	4	BG951790 MR1-CT073
35	443	19.9	481	4	BM62514 UI-E-E01-
36	437.4	19.6	531	4	BM685396 UI-E-E01-
37	427.4	19.2	484	7	N31453 yx55d08.r1
38	422	18.9	448	5	BT733195 UI-E-E01-
39	405.4	18.2	419	4	BM718146 UI-E-E01-
40	402.6	18.1	792	7	CK470840 AGENCOURT
41	401.4	18.0	508	2	BE235901 145906 MA
42	391	17.5	766	2	BE275846 601121656
43	388	17.4	713	6	BY714331 BY714331
44	385.2	17.3	862	7	CK469826 AGENCOURT
45	379.8	17.0	1067	4	BM562042 AGENCOURT

ALIGNMENTS

RESULT 1
B0879880
LOCUS
DEFINITION
AGENCOURT_8241531_lupsk1_dorsal_root_ganglion Homo sapiens cDNA
clone IMAGE:6180559 5', mRNA sequence.
B0879880
B0879880.1 GI:22271888
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Cranial; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
NIH-MGC http://mgc.ncl.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs@ncl.nih.gov
Tissue Procurement: Dr. James R. Lupski
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov
Plate: LNLN1363 row: 1 column: 08
High quality sequence stop: 623.
Location/Qualifiers
1. 963
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6180559"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/lab_1id="lupsk1_dorsal_root_ganglion"
/note="Vector: pCMV-Sport6 (Life Technologies); Site_1:
NotI; Site_2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACGACGCGCGCG-3' and
5'-GACGATCTGATCGGACGCGCGCGCTT(15)-3'. Size selected >
1 kb for average insert length 1.7 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life


```

QY 1050 CAGGTGTCCTGTTGGCTGTGATGATGACACGAGACTGCTCGATGAGCCGAGC 1109
D 481 CAGGTGTCCTGTTGGCTGTGATGATGACACGAGACTGCTCGATGAGCCGAGC 540
QY 1110 ATGTACTGGAATTAAGCCCGAGCCACAGCCCTTACACAGCTGCTTCCGCCAGTTCCT 1169
D 541 ATGTACTGGAATTAAGCCCGAGCCACAGCCCTTACACAGCTGCTTCCGCCAGTTCCT 600
QY 1170 CGAGCTCTCTTTAAGGGCTCCACCTTCAACATCAGCTCGAACAAGAGAGATGAGTTC 1229
D 601 CGAGCTCTCTTTAAGGGCTCCACCTTCAACATCAGCTCGAACAAGAGAGATGAGTTC 660
QY 1230 CAGCCCATCAGAGAGACGAGAGAGATGCTCAGCTGATCATTTGGCCCTTCTAGGC 1289
D 661 CAGCCCATCAGAGAGACGAGAGAGATGCTCAGCTGATCATTTGGCCCTTCTAGGC 720
QY 1290 CTGCAATGCTCCATGATCACCATCTCTCCAGGC--CAACTCAGAGACCAACTACTGTGGC 1347
D 721 CTGCAATGCTCCATGATCACCATCTCTCCAGGC--CAACTCAGAGACCAACTACTGTGGC 780
QY 1348 CCAAGAGGGA---TCCCTTCTCAGAGGGCTGCCCAAAACCAACAGAGCAGCAAC 1404
D 781 CCAAGAGGGAATTCCTTCTCAGAGGGCTGCCCAAAACCAACAGAGCAGCC 840
QY 1405 AGAAGCTTACGGGCGCCAGAGAACAA 1429
D 841 AACCAGAACTTTTACGGGCGCCCA 865

RESULT 3
AK006549
LOCUS 1699 bp mRNA linear HTC 03-APR-2004
DEFINITION Mus musculus adult male testis cDNA, RIKEN full-length enriched
library, clone:1700030H21 product:Vitellogenin macular dystrophy 2
homolog (human), full insert sequence.
ACCESSION AK006549
VERSION AK006549.1 GI:12839710
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636

REFERENCE
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159

REFERENCE
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kikunishi, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kaishogi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matshiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuda, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861

REFERENCE
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
PANTOM Consortium.

```

```

TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5
AUTHORS The PANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6 (bases 1 to 1699)
AUTHORS Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H.,
Arai, K., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,
Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,
Hiraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,
Kaishogi, K., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K.,
Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,
Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y.,
Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T.,
Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K.,
Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
TITLE Direct Submission
JOURNAL Submitted (10-JUL-2000) Yoshihide Hayashizaki. The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gscc.riken.jp,
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
COMMENT Please visit our web site (http://genome.gsc.riken.jp/) for further
details.
FEATURES
SOURCE 1..1699
location=Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:1700030H21"
/db_xref="taxon:10090"
/clone="1700030H21"
/sex="male"
/tissue_type="testis"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
misc_feature 1..1699
/note="Vitellogenin macular dystrophy 2 homolog (human)"
(MGI|MGI:1346332)

ORIGIN
Query Match 33.8%; Score 754.2; DB 3; Length 1699;
Best Local Similarity 77.3%; Pred. No. 1,66-179;
Matches 983; Conservative 0; Mismatches 248; Indels 40; Gaps 4;

QY 273 GAACACAGCTGATGTTGAGAACTGACTCTGTATTGCGACAGCTACATCCAGCTCATC 332
D 1 GATCAGCAGCTGTTGTTGAGAACTGCTGTGATGCGACAGCTACATCCAGCTCATC 60
QY 333 CCATTCCTTCGCGCGGGCTCTCTACGACCGTGTGACCGCGCTGTGGAACAG 392
D 61 CTTATATCTTCCTTCGCGGGCTCTCTATGTTACATTTGTTGAGACCGCGCTGTGGAACAG 120
QY 393 TAGAGAACTGCGGTGGCCGACCGCTCATGAGCTGTGTCGGGCTTCTGGAAGGC 452

```

Db	121	TACGAAACTTGCCTTGGCCCGACCGCTCTCATATCCAGGTGTACGCTTGTGGTAGAGGC	180
Oy	453	AAGACGACGAGGCGCGCTGTCTGCGAGCGACGCTCATCCGCTACGCCAACCTGGCAAC	512
Db	181	AAGGATAGAGAGGCGTTTGCTGCGGCCAGGCTCATCCGCTAGGCCATCTTGGGCGAA	240
Oy	513	GTGCTCATCCCTGGCGAGCGCTGACGACCCGACGCTCTCAAGCCCTTCCCGAGCCACAGC	572
Db	241	GTGCTCATCCCTGGCGAGCATCAGCACCTTGCTCTCAAGCCCTTCCCATCTTCAACAC	300
Oy	573	CTGTGTGACGAGGCTTTATGATCCCGCAGAAACAAGCAGTGTGGAAATGAGACTTA	632
Db	301	CTGTGTGTGACGAGTCTTTATGACCCATGGGGGAACAAGCATGTTGCGAAAGTTGGGCTTA	360
Oy	633	CCACACAAATGTTCTGGGTGCCCCGGGTGTGTGTTTGGCAACTGTCAATGAAGCGTGC	692
Db	361	CCACA-ACACATTTCTGGGTGCCCCGGGTGTGTGTTTGGCAACTGTCTCAATGAAGGCTTAT	419
Oy	693	CTTGAAGTTCGATCCGGAACCTTATCTGTCTCAGAGCCGTGCGAAACGAGATGAACCC	752
Db	420	CTTGAAGTTCGATCCGGAACCCGTCTGTCTCAGAGCCGTGCGAAACGAGATGAAGTGTACT	479
Oy	753	TTGCGTACTCAGGTGGGACACCTGTATAGCTTACGACTGTAAATGATATCCCATGTGTAT	812
Db	480	TTGCGTACTCAGGTGGGACAGCTGTATAGCTTACGACTGTAAATGATATCCCATGTGTATAC	539
Oy	813	ACACAGGTGTGACTGTGGGGGTGTACGCTTCTGTCTGACCTTGTCTAGTTGGCGCGAC	872
Db	540	ACACAGGTGTGTACGAGTGGCAGTATACGCTTCTTCTTGCATGCTTATCGGAGAGGAC	599
Oy	873	TTTCTGAACCCAGCGACAGGCTTACCTGGCATAGACTTGACCTGATTTGTGACGAAATTTGGC	932
Db	600	TTTCTGAACCCAGCGACAGGACTTACCCAGGCGATGAGATGATCTGTTGTGCTGTCTTC	659
Oy	933	ACGTTCCGACGTTCTTCTTATATGTGTGCTGCGTGAAGGTGGGAGAGCACTCATTAAC	992
Db	660	ACATCTCTGCAATTTCTTATTTCTACATGGGCTGGCTGAAGGTGGGAGAGCACTCATTAAC	719
Oy	993	CCCTTGAAGAGGATGATGATTTTGAACCAATGATTTGTGACAGGAATTTGGC	1052
Db	720	CCCTTGGGAGGAGCATGATATATTTTGAAGTAACTGATCAATTTGAACAAGAACTGTGAC	779
Oy	1053	GTGTCCCTGTGTGTGTGATGAGATGACCAAGGACCTGCTGCGATGTGAGCCGACATG	1112
Db	780	GTGTCCCTGTGTGTGTGATGAGATGACCAAGGACCTGCTGCGATGTGAGCCGACATG	839
Oy	1113	TACTGGAATAACCCGAGCCACAGCCCTCTTACAGACTGTTCCGCCAGTTCGGTGTGA	1172
Db	840	TACTGGAACGAGGACAGCGCTGTACGCCCTCTTACAGACTGTTCCGCCAGTTCGGTGTGA	1172
Oy	1173	GCCTCTTATATGGGCTTCACCTTCAACATGACGCTGAAACAAGAGAGATGAGTTTCAG	1232
Db	900	CATTTCTTTCATGGGCTTCACCTTCAACATGACGCTTCAAGCTTAAAGAAACATCTTAAGCTTTGG	959
Oy	1233	CCCATTCAGGAGG-----ACGAGAGAGATGCTACGCTGGGCTCATTTGGCGCG	1280
Db	960	TCAAAAGGAGGAGGCTGACACGATATGAAGAAAGAGAGTGCGCTATAGGACAGCACTAGGCTGCG	1019
Oy	1281	TTTCTTAGGCGCTGACGATCCCATGATTCACATCCCTCCAGGCGAACTCAAGACCAACTA	1340
Db	1020	TTTCTTAGGAGTGCAGACCCAAAACCTACATCTTCCCTTGAAGAGCTTAAAGACCAACTA	1079
Oy	1341	CTGTGGGCCAAGAGGAGATCTTCTTCAACGAGGCGTGGCCCAAAAACAAGAGGAGCC	1400
Db	1080	TTGTGTCTTAAAG-----ACCCCTCTCTCGAAGGCGAGTGTATAGAGTCCCAACAGAAA	1133
Oy	1401	AAACAGAACTTATGGGCGCAGGAAGACAAAGGCTGTGAAGCTTAAAGCTGTGAGCGCC	1460
Db	1134	AACCAAGAAAG-----ATGTCTGAAATTTAAAGGAGGCTGTGAGCTTC	1172
Oy	1461	TTCAATGTGTGCGCCACTGTATCAGAGCGAGGCTTACTACATGTGCCACAGAGCGCCCTTC	1520

Db	1173	TTGAAATGTTTCCAAAGGTTTAAGAGGAGAGCTCCATTGTGGCCACAGGCACTCAGC	1233
Oy	1521	AGCCCCACTCC	1531
Db	1233	AGCCACCTTAC	1243

LOCUS	BX055540	711 bp	mRNA	linear	EST 04-FEB-2003
DEFINITION	BX055540.0 Soares melanocyte 2b2bHm Homo sapiens cDNA clone				
ACCESSION	U009988B1582 ; IMAGE:265647				
VERSION	BX055540				
KEYWORDS	BX055540.1 GI:27827758				
SOURCE	EST.				
ORGANISM	Homo sapiens (human)				
	Homo sapiens				

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1	Mammalia; Eutheria; Primates; Carnivora; Eutelestomii (basses 1 to 71)	Ebert, L., Hell, O., Hennig, S., Neubert, P., Patsch, E., Peters, M., Rabelof, U., Schneider, D. and Korn, B.	Human Unigeneset - R2PD3	Unpublished (2003)
			Contact: Ina Rolf	

Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
 RZPDLib; I.M.A.G.E. CDNA Clone Collection;
 Human UniGeneset - RZPD3 (RZPDLIB No. 972)
[http://www.rzpd.de/CloneCards/cgi-
 bin/showlib.pl.cgi?response=libNo=972](http://www.rzpd.de/CloneCards/cgi-bin/showlib.pl.cgi?response=libNo=972) Contact: Ina Rolfs
 RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
 Heubnerweg 6, D-14059 Berlin, Germany
 Tel.: +49 30 32639 110
 Fax: +49 30 32639 111
www.rzpd.de
 This clone is available royalty-free from RZPD;
 contact RZPD (clone@rzpd.de) for further information. Seq primer
 M13r, Primer sequence: TTTCACGAGAAACGATATAC.
 Location/Qualifiers
 1..711

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGP998B16582 ; IMAGE:265647"
/gen="Male"
/tissue_type="melanocyte"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Scotars melanocyte 2NDHM"
/vector="pTRT3D (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dAT) primer 15',
TGTTACCACTCGAAGTCGACGACGCCGATTTTCTTTTCTTTT 3',
the cDNA-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pTRT3 vector
(Pharmacia). Library constructed by Benito Soares and
M.Fátima Bonaldi. RNA from normal foreshkin melanocytes
(FS374) was kindly provided by Dr. Anthony P. Albino."

```

Query Match	Score	DB 5;	Length
30.6%	681.8;	711;	
Best Local Similarity	99.6%		

```
conservative    0; Mismatches   2; Indels     1; Gaps      1;
```

[illegible]

1392 AAGCAGCCAAACGAACTTAGGGGCCAGGAAGACAAGGCCTGGAAGCTTAAGCTT 1451
61 AAGGTGGCCATTATCAACTTT

QY 1452 GTGAGCGCTTCAAGTCTGGCCCACTGATCAGAGGCCAGGCTACTACAGTGGCCCAAG 1511
 DB 121 GTGAGCGCTTCAAGTCTGGCCCACTGATCAGAGGCCAGGCTACTACAGTGGCCCAAG 180
 QY 1512 ACGCCCTCAGGCCCACTCCCATGTTCTTCCCTAGAACCATCAGCGCCGTCAAGCTT 1571
 DB 181 ACGCCCTCAGGCCCACTCCCATGTTCTTCCCTAGAACCATCAGCGCCGTCAAGCTT 240
 QY 1572 CACAGTGTACAGGCATAGACACCAAGACAAAGCTTAAAGACTGTGATTTCTGGGGCC 1631
 DB 241 CACAGTGTACAGGCATAGACACCAAGACAAAGCTTAAAGACTGTGATTTCTGGGGCC 300
 QY 1632 AAGAAAGTTTGAATGCTCTCAGAGAGGATGGGGCTTGATGAGACCCAGAAAGTA 1691
 DB 301 AAGAAAGTTTGAATGCTCTCAGAGAGGATGGGGCTTGATGAGACCCAGAAAGTA 360
 QY 1692 TCTCAAGTGAAGAGAAACCTGTGAGTTTAACTGACGATATGCGAGATCCCGAA 1751
 DB 361 TCTCAAGTGAAGAGAAACCTGTGAGTTTAACTGACGATATGCGAGATCCCGAA 420
 QY 1752 AATCACTCAAGAAACCTTTGGAAACATCACCACCAACATACACACTCAAGAT 1811
 DB 421 AATCACTCAAGAAACCTTTGGAAACATCACCACCAACATACACACTCAAGAT 480
 QY 1812 CACATGATCCTTATTTGGGCGCTTGAAACAGGGATGAAGACATTCCTAACCTGCTTCC 1871
 DB 481 CACATGATCCTTATTTGGGCGCTTGAAACAGGGATGAAGACATTCCTAACCTGCTTCC 540
 QY 1872 TAATGGGAGATGCTTCGCGACGAGATCCTCACTGTGTATCACCAGCAGCACTGATC 1931
 DB 541 TAATGGGAGATGCTTCGCGACGAGATCCTCACTGTGTATCACCAGCAGCACTGATC 600
 QY 1932 CAGTCAACGACCATACAGTGTTCACACTGAAGAACGTGTCTTCAACACGCTGAATCAA 1991
 DB 601 CAGTCAACGACCATACAGTGTTCACACTGAAGAACGTGTCTTCAACACGCTGAATCAA 660
 QY 1992 TGGTACCTTAATAGATAAATCCAGACTACTTCA 2028
 DB 661 TGGTACCTTAATAGATAAATCCAGACTACTTCA 697

RESULT 5
 LOCUS BU741926/c 676 bp mRNA linear EST 10-OCT-2002
 DEFINITION UI-E-BO1-aj-a-j-23-0-UI.s1 UI-E-BO1 Homo sapiens cDNA clone
 ACCESSION BU741926
 VERSION BU741926.1 GI:23687736
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 Bernaldo,M.F., Lennon,G. and Soares,M.B.
 Normalization and subtraction: two approaches to facilitate gene
 discovery
 Genome Res. 6 (9), 791-806 (1996)
 JOURNAL MEDLINE 97044477
 PUBMED 8889548
 COMMENT Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@uiowa.edu
 Tissue Procurement: Dr. Gregg Hageman
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research

Genetics (www.iesgen.com).
 Seq primer: M13 FORWARD
 PolyA=Yes.
 Location/Qualifiers
 1..676
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-E-BO1-aj-a-j-23-0-UI"
 /tissue_type="fetal eye"
 /dev_stage="fetal"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /clone_id="UI-E-BO1"
 /notes="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
 modified polylinker; Site 1: Score 1; Site 2: Not I;
 UI-E-BO1 is a normalized cDNA library containing the
 following tissue(s): fetal eye. The library was
 constructed according to Bernaldo, Lennon and Soares,
 Genome Research, 6:791-806, 1996. First strand cDNA
 synthesis was primed with an oligo-dT primer containing a
 Not I site. Double stranded cDNA was ligated to an EcoR I
 adaptor, digested with Not I, and cloned directionally
 into pT73-Pac vector. The oligonucleotide used to prime
 the synthesis of first-strand cDNA contains a library tag
 sequence that is located between the Not I site and the
 (dT)18 tail. The sequence tag for this library is
 GCGGTATACC. This library was created for the program, Gene
 Discovery in the Visual System, supported by National Eye
 Institute (NEI).
 TAG_TISSUE=human fetal eye
 TAG_LIB=UI-E-BO1
 TAG_SEQ=CCGATATACC"

ORIGIN

Query Match 29.5%; Score 657; DB 5; Length 676;

Best Local Similarity 99.1%; Pred. No. 5.6e-155;

Matches 671; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 1543 CCTTGAACCATTCAGCGCCGCTCAAGCTTCACTGTCACAGCATAGACCAAGACA 1602
 DB 676 CCTTGAACCATTCAGCGCCGCTCAAGCTTCACTGTCACAGCATAGACCAAGACA 618
 QY 1603 AAAGCTTAAAGACCTGTAGTTCTGGGACCAAGAAAGTTTGAATGCTCTCAGAGACG 1662
 DB 617 AAAGCTTAAAGACCTGTAGTTCTGGGACCAAGAAAGTTTGAATGCTCTCAGAGACG 558
 QY 1663 ATGGGGCTTGATGAGACACCCAGAGATATCTCAAGTGAAGAGAAACCTGTGAAGTTTA 1722
 DB 557 ATGGGGCTTGATGAGACACCCAGAGATATCTCAAGTGAAGAGAAACCTGTGAAGTTTA 498
 QY 1723 ACCTGACGGATATGCCAGAGATCCCCGAAATATCACTTCAAGAAACCTTTGGAACATAC 1782
 DB 497 ACCTGACGGATATGCCAGAGATCCCCGAAATATCACTTCAAGAAACCTTTGGAACATAC 438
 QY 1783 CAACCAACATACACATACATCACTCAAGATCATGATATCTTATTTGGGCTTTGAAAAACA 1842
 DB 437 CAACCAACATACACATACATCACTCAAGATCATGATATCTTATTTGGGCTTTGAAAAACA 378
 QY 1843 GGGATGAAGCATTTCTTAACCTTGCTTCTTAATGGGAGATGCTTGCCAGCAGGTCTCA 1902
 DB 377 GGGATGAAGCATTTCTTAACCTTGCTTCTTAATGGGAGATGCTTGCCAGCAGGTCTCA 318
 QY 1903 CTTGTGTATACCAAGACAGACATCTGATCCAGTCAAGCCATACAGCTGTCCACATGAA 1962
 DB 317 CTTGTGTATACCAAGACAGACATCTGATCCAGTCAAGCCATACAGCTGTCCACATGAA 258
 QY 1963 GAACGTGTCTCAACAAGACCTGATCAATGATGCTTATAGATTAATAAATCCAGACT 2022
 DB 257 GAACGTGTCTCAACAAGACCTGATCAATGATGCTTATAGATTAATAAATCCAGACT 198
 QY 2023 ACTTACGCTTTAATGCTTTTATTCATTAATAAAGCTGAAAGCTAGACTGAACATTTGA 2082
 DB 197 ACTTACGCTTTAATGCTTTTATTCATTAATAAAGCTGAAAGCTAGACTGAACATTTGA 138

QY 2083 AACATTAACTGAGTCTGATTCAGAGTCGGGAACCCCTTATGTTCTATCTGAATCCAG 2142
Db 137 AACATTAACTGAGTCTGATTCAGAGTCGGGAACCCCTTATGTTCTATCTGAATCCAG 78
QY 2143 ACAGCCACACCTTATGATCTGCTCCAACTAATGATTAATTAATCAATCTGCTTA 2202
Db 77 ACAGCCACACCTTATGATCTGCTCCAACTAATGATTAATTAATCAATCTGCTTA 18
QY 2203 AAAAAAAAAAAAAAAAAA 2219
Db 17 AAAAAAAAAAAAAAAAAA 1

RESULT 6 659 bp mRNA 1linear EST 09-OCT-2002
LOCUS BU730894/c
DEFINITION UI-E-C11-afw-i-11-0-UI s1 UI-E-C11 Homo sapiens cDNA clone
ACCESSION BU730894
VERSION BU730894.1 GI:23655244
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 659)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
8889548

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@iowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 FORWARD
POLY-A=yes

FEATURES
source
Location/Qualifiers
1..659
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-C11-afw-i-11-0-UI"
/tissue_type="RPE and Choroid"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_id="UI-E-C11"
/note="Organ: eye; Vector: pRTT3-Pac (pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-E-C11 is a normalized cDNA library containing the
following tissue(s): RPE and Choroid. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pRTT3-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
(d18 tail). The sequence tag for this library is ACCTA.
This library was created for the program, Gene Discovery
in the Visual System, supported by National Eye Institute

ORIGIN
(NEI).
TAG_TISSUE=RPE and Choroid
TAG_LIB=UI-E-C11
TAG_SEQ=ACCTA"

Query Match 29.2%; Score 651.6; DB 5; Length 659;
Best Local Similarity 99.2%; Pred. No. 1.3e-153;
Matches 654; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1561 CCTCAAGCTTCAAGTGTCAAGAGCAATAGACCAAGCAAAAGCTTAAAGCTGTGA 1620
Db 659 CGTCAAAGCTTCAAGTGTCAAGAGCAATAGACCAAGCAAAAGCTTAAAGCTGTGA 600
QY 1621 GTTCTGGGGCCCAAAAAGTTTGAATTCCTCTCAAGAGCGATGCGCCCTTGATGAGC 1680
Db 599 GTTCTGGGGCCCAAAAAGTTTGAATTCCTCTCAAGAGCGATGCGCCCTTGATGAGC 540
QY 1681 ACCCAAGATATCTCAAGTGAAGGAAAAGTGTGAGTTTAACTGAGGATATGCCAG 1740
Db 539 ACCCAAGATATCTCAAGTGAAGGAAAAGTGTGAGTTTAACTGAGGATATGCCAG 480
QY 1741 AGATCCCGGAAAATCACTTGAAGAACTTTGGAACATACCAACCAATACACACTA 1800
Db 479 AGATCCCGGAAAATCACTTGAAGAACTTTGGAACATACCAACCAATACACACTA 420
QY 1801 CACTCAAAAGTACATGATATCTTATTTGGGCTTGGAAAACGGGATGAAGACATTTCT 1860
Db 419 CACTCAAAAGTACATGATATCTTATTTGGGCTTGGAAAACGGGATGAAGACATTTCT 360
QY 1861 AACCTGCTTCTAATGAGGATGCTTGGCCAGGCTCTCACTGTGTGACACACAGA 1920
Db 359 AACCTGCTTCTAATGAGGATGCTTGGCCAGGCTCTCACTGTGTGACACACAGA 300
QY 1921 GGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1980
Db 299 GGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
QY 1981 CCTGAATCAATGTTAGCTTAATGATGATGATGATGATGATGATGATGATGATGAT 2040
Db 239 CCTGAATCAATGTTAGCTTAATGATGATGATGATGATGATGATGATGATGATGAT 180
QY 2041 TTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2100
Db 179 TTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
QY 2101 TGGATTCAGAGTCGGGAACCCCTTATGTTCTATCTGAATCCAAAGCAGCACTTATGAT 2160
Db 119 TGGATTCAGAGTCGGGAACCCCTTATGTTCTATCTGAATCCAAAGCAGCACTTATGAT 60
QY 2161 ACTGCCCAACTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2219
Db 59 ACTGCCCAACTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1
RESULT 7 674 bp mRNA 1linear EST 09-OCT-2002
LOCUS BU731809/c
DEFINITION UI-E-C11-afw-i-11-0-UI s1 UI-E-C11 Homo sapiens cDNA clone
ACCESSION BU731809
VERSION BU731809.1 GI:23657073
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 674)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477

PUBMED 8889548
 COMMENT Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@uiowa.edu
 Tissue Procurement: Dr. Gregg Hageman
 CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).
 Seq primer: M13 FORWARD
 PolyA+yes.

FEATURES
 source Location/Qualifiers
 1..674
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-E-C11-af1-g-11-0-UI"
 /rname_type="RPF and Choroid"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /clone_lib="UI-E-C11"
 /note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site_1: EcoR I; Site_2: Not I; UI-E-C11 is a normalized cDNA library containing the following tissue(s): RPF and Choroid. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is ACCTA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI).
 TAG_TISSUE=RPE and Choroid
 TAG_LIB=UI-E-C11
 TAG_SEQ=ACCTA"

ORIGIN

Query Match 29.1%; Score 648.8; DB 5; Length 674;
 Best Local Similarity 97.9%; Pred. No. 6.7e-153;
 Matches 656; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

1548 GAACCATCAGGCGCGTCAAGCTTCAAGTGTCAAGGCATAGACACCAAGACAAAAGC 1607
 670 GCACGAGGAGGCGCGTCAAGCTTCAAGTGTCAAGGCATAGACACCAAGACAAAAGC 611
 1608 TTAAGAGCTGTGAGTTCCTGGGCGCAAGAAAGTTTGAAATGCTCTCAAGAGAGGATGGG 1667
 610 TTAAGAGCTGTGAGTTCCTGGGCGCAAGAAAGTTTGAAATGCTCTCAAGAGAGTGGG 551
 1668 GCCTTGATGAGCAAGCAAGATATCTCAAGTGAAGAGAAACGTGTGAGTTTAACTTG 1727
 550 GCCTTGATGAGCAAGCAAGATATCTCAAGTGAAGAGAAACCGTGTGAGTTTAACTTG 491
 1728 ACGGATATGCCAGAGATCCCGAAATATCACTCAAGAACTTTGGAACAATCAACAAC 1787
 490 ACGGATATGCCAGAGATCCCGAAATATCACTCAAGAACTTTGGAACAATCAACAAC 431
 1788 AAGTACACACTAATCAAGATCAAGTGAATCTTATTTGGGCTTGGAAAAAGAGGAT 1847
 430 AAGTACACACTAATCAAGATCAAGTGAATCTTATTTGGGCTTGGAAAAAGAGGAT 371
 1848 GAAGACACTTCTTAACCTGTTCTTAATGGGATGCTTCGACAGCAGAGTCTCAACCTGT 1907

Db 370 GAAGACACTTCTTAACCTGTTCTTAATGGGATGCTTCGACAGCAGAGTCTCAACCTGT 311
 Qy 1908 GTGTACACCGAGCGGAGACATGATCCAGTCAAGCATACAGCTGTCCAGACTGAAGACG 1967
 Db 310 GTGTACACCGAGCGGAGACATGATCCAGTCAAGCATACAGCTGTCCAGACTGAAGAC 251
 Qy 1968 TGTCTCTCAACAGCCTGAATCAAAATGTTAGCTTAATAGATAAAATCCAGACTACTTC 2027
 Db 250 TGTCTCTCAACAGCCTGAATCAAAATGTTAGCTTAATAGATAAAATCCAGACTACTTC 191
 Qy 2028 AGCCTTAAATGCTTTTATTCATTAATAAACTGTGAAGCTAGACTGAACATTGGAAACAT 2087
 Db 190 AGCCTTAAATGCTTTTATTCATTAATAAACTGTGAAGCTAGACTGAACATTGGAAACAT 131
 Qy 2088 TTAACCTGAGACTGTGATTAAGATGAGAGTGGGAAACCTTGATCTATCTGAATCAAGACAGC 2147
 Db 130 TTAACCTGAGACTGTGATTAAGATGAGAGTGGGAAACCTTGATCTATCTGAATCAAGACAGC 71
 Qy 2148 CACACCTTAGTACTGCTCCCAAACTAATGAGTTAATTAATACAAATACCTGTTAAAAA 2207
 Db 70 CACACCTTAGTACTGCTCCCAAACTAATGAGTTAATTAATACAAATACCTGTTAAAAA 11
 Qy 2208 AAAAAAAAAA 2217
 Db 10 AAAAAAAAAA 1

RESULT 8
 CO396179 756 bp mRNA linear EST 01-JUL-2004
 LOCUS AGENCOURT 27872871 NIH MGC 212 Homo sapiens cDNA clone
 DEFINITION IMAGE:30923539 5', mRNA sequence.
 ACCESSION CO396179
 VERSION CO396179.1 GI:49578095
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 756)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cgabds-remail.nih.gov
 Tissue Procurement: Mary Hendrix
 CDNA Library Preparation: M. Bento Soares, University of Iowa
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LMLN)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LMLN at:
 http://image.llnl.gov
 Plate: NDAM164 row: 1 column: 20
 High quality sequence. Stop: 571.

FEATURES

source

Location/Qualifiers
 1..756
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30923539"
 /rname_type="Chondrosarcoma Lung Metastasis cell lines"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH MGC 212"
 /note="Organ: Lung; Vector: pTX-Aac; Site_1: EcoR I; Site_2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated

with Sc0r I adaptor, digested with Not I and then cloned directionally into pUX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is GATAGGCCCA. Tissue was provided by Mary Hendrix."

ORIGIN

Query Match 28.8%; Score 641.8; DB 7; Length 756;
Best Local Similarity 98.8%; Pred. No. 4e-151;
Matches 646; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```
QY 417 CGCCTCATGAGCTGGGTGTCGGGCTTCGTCGCAAGCAGACGACCAAGCCGCGTCTG 476
DB 8 CGCCTCATGAGCTGGGTGTCGGGCTTCGTCGCAAGCAGACGACCAAGCCGCGTCTG 476
QY 477 CGCGGACGCTCATGTCGCTAGCGCAACCTGGGCAACGTCATCTCGGCGAGGCTGAC 536
DB 68 CGCGGACGCTCATGTCGCTAGCGCAACCTGGGCAACGTCATCTCGGCGAGGCTGAC 536
QY 537 ACCGAGTCTACAGGCTTCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 596
DB 128 ACCGAGTCTACAGGCTTCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 596
QY 597 CGCGGACGCTCATGTCGCTAGCGCAACCTGGGCAACGTCATCTCGGCGAGGCTGAC 656
DB 188 CGCGGACGCTCATGTCGCTAGCGCAACCTGGGCAACGTCATCTCGGCGAGGCTGAC 656
QY 657 TGGGTGTGTTGGCAACCTGTCATAGAGCGCTGGGCTGGAGGTCGAAATCCGGGACCT 716
DB 248 TGGGTGTGTTGGCAACCTGTCATAGAGCGCTGGGCTGGAGGTCGAAATCCGGGACCT 716
QY 717 ATCTGCTCCAGAGCTGCTGTAACGAGATGATACACCTTGGGCTGAGTGTGAGGCTG 776
DB 308 ATCTGCTCCAGAGCTGCTGTAACGAGATGATACACCTTGGGCTGAGTGTGAGGCTG 776
QY 777 TATGCTTACAGCTGATAGTATCCAGCTGCTGTAACGAGTGTGAGGCTGAGGCTG 836
DB 368 TATGCTTACAGCTGATAGTATCCAGCTGCTGTAACGAGTGTGAGGCTGAGGCTG 836
QY 837 TACAGCTTCTTCTGCTGATGTTGGGCGGCGGAGTTCGTAACCGGCGGCGGCTGAC 896
DB 428 TACAGCTTCTTCTGCTGATGTTGGGCGGCGGAGTTCGTAACCGGCGGCGGCTGAC 896
QY 897 CCTGCGCATGAGCTGAGCTGCTGTCGTCGCTGCTGCTGCTGCTGCTGCTGCTGCT 956
DB 488 CCTGCGCATGAGCTGAGCTGCTGTCGTCGCTGCTGCTGCTGCTGCTGCTGCTGCT 956
QY 957 GTTGGCTGCTGTAAGGCTGAGGAGGCTGATCAACCCCTTGGAGAGGATGATGAT 1016
DB 548 GTTGGCTGCTGTAAGGCTGAGGAGGCTGATCAACCCCTTGGAGAGGATGATGAT 1016
QY 1017 TTTGAGACCAACTGATGTCGACAGGAAATTTGAGGCTGCTGCTGCTGCTGCTG 1070
DB 608 TTTGAGACCAACTGATGTCGACAGGAAATTTGAGGCTGCTGCTGCTGCTGCTG 1070
```

RESULT 9

LOCUS B1480798

DEFINITION

H2RPE-0430 Human Retinal Pigment Epithelium (2) Homo sapiens cDNA
5' similar to vitelliform macular dystrophy (Best disease, mRNA
sequence.

ACCESSION

VERSION B1480798

KEYWORDS

EST.

SOURCE

ORGANISM Homo sapiens (human)

REFERENCE

1 (bases 1 to 666)

AUTHORS

Buraczynska, M., Weisz, A., J., Zarepari, S., Farjo, R., Filipova, E.,
Yuan, Y., MacNee, S. P., Hughes, B., and Swaroop, A.

TITLE

Towards an expression profile of native human retinal pigment
epithelium: Identification of a non-redundant set of more than 1100

JOURNAL

Unpublished (2001)

COMMENT

Contact: Swaroop, A.
Department of Ophthalmology and Visual Sciences
Kellogg Eye Center, University of Michigan
540 KEC, 1000 Mail St., Ann Arbor, MI 48105, USA
Tel: 734 615 2246
Fax: 734 647 0228
Email: swaroop@umich.edu

FEATURES

source

ORIGIN

Query Match 28.7%; Score 640.4; DB 4; Length 666;
Best Local Similarity 99.4%; Pred. No. 8.9e-151;
Matches 652; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

```
QY 10 CCACGAGCTTATGCGCAGACCTTCGTTGGGATCATCGACCCACCTGGAACCCACCTG 69
DB 12 CCACGAGCTTATGCGCAGACCTTCGTTGGGATCATCGACCCACCTGGAACCCACCTG 69
QY 70 ACCGAGCCCACTGCTGAGCGGCACTGCTGCGCATGACCATCATCTTACACAGCCAG 129
DB 72 ACCGAGCCCACTGCTGAGCGGCACTGCTGCGCATGACCATCATCTTACACAGCCAG 129
QY 130 TGGCTAATGCGCGCTTATGAGCTCTTCTGCGGCTGCTGCTGCTGCTGCTGCTGCT 189
DB 132 TGGCTAATGCGCGCTTATGAGCTCTTCTGCGGCTGCTGCTGCTGCTGCTGCTGCT 189
QY 190 ACAAGCTGTAATGCGAGGTTCTTAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 249
DB 192 ACAAGCTGTAATGCGAGGTTCTTAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 249
QY 250 TTTATAGGCTGCGCCTACAGGAGAGAAACAAGCTGATGTTTGAAGACTGACTGTAT 309
DB 252 TTTATAGGCTGCGCCTACAGGAGAGAAACAAGCTGATGTTTGAAGACTGACTGTAT 309
QY 310 GCGACAGCTACATCACTGATCCCATTTCTTCTGCTGCGGCTTCTACGAGGCTG 369
DB 312 GCGACAGCTACATCACTGATCCCATTTCTTCTGCTGCGGCTTCTACGAGGCTG 369
QY 370 TCGTGAACCGCTGTTGGAACCAAGTACAGACCTTGGCGGCTGCGGCTGATGAGCC 429
DB 372 TCGTGAACCGCTGTTGGAACCAAGTACAGACCTTGGCGGCTGCGGCTGATGAGCC 429
QY 430 TGGTGTGCGGCTTCTGTCGAGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 489
DB 432 TGGTGTGCGGCTTCTGTCGAGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 489
QY 490 TCCGCTAAGCGCAACCTGGGCAACGTCATCTCTGCGAGGCTGAGACCGCACTTACA 549
DB 492 TCCGCTAAGCGCAACCTGGGCAACGTCATCTCTGCGAGGCTGAGACCGCACTTACA 549
QY 550 AGGCGTTCGCCGCGCGGCGGAGCACTGTCGCAAGAGGCTTTATGATCCGCGAGAAC 609
DB 552 AGGCGTTCGCCGCGCGGCGGAGCACTGTCGCAAGAGGCTTTATGATCCGCGAGAAC 609
QY 610 ACCAGTTGAGAGAAATGAGCTTACCAACAACATGTTTGGGTCCTGCGGCTGAGTGG 665
DB 612 ACCAGTTGAGAGAAATGAGCTTACCAACAACATGTTTGGGTCCTGCGGCTGAGTGG 665
```

RESULT 10	
CA389968	639 bp mRNA linear EST 06-NOV-2002
LOCUS	ca810403.y1 Human Retinal pigment epithelium/choroid cDNA
DEFINITION	(Un-normalized, unamplified): cs Homo sapiens CDNA clone ca810403 5', mRNA sequence.
ACCESSION	CA389968
VERSION	CA389968.1 GI:24720628
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 639)
AUTHORS	Wiatow, G., Bernstein, S.L., Wyatt, M.K., Farris, R.N., Behal, A., Touchman, J.W., Bouffard, G., Smith, D. and Peterson, K. Expressed sequence tag analysis of human RPE/choroid for the NEIBank Project: Over 6000 non-redundant transcripts, novel genes and splice variants
TITLE	Mol. Vis. 8 (4), 205-220 (2002)
JOURNAL	22103460
MEDLINE	12107410
PUBMED	Contact: Wiatow G
COMMENT	Section on Molecular Structure and Function National Eye Institute 6/331, NIH, Bethesda, MD 20892-2740, USA Tel: 301 402 3452 Fax: 301 496 0078 Email: graeme@helix.nih.gov Plate: 104 row: h column: 03 Seq primer: M13RPI reverse primer (ABI). Location/Qualifiers 1..639 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="ca810403" /tissue_type="RPE/choroid" /dev_stage="Adult" /lab_host="EMD10B" /clone_lib="Human Retinal pigment epithelium/choroid cDNA (Un-normalized, unamplified): cs" /note="Organ: Eye; Vector: pCMVSPORT6; Two different donor eyes (75-80 years old) yielded approximately 600 mg of dissected RPE/choroid tissue. This in turn yielded 340 ug of total RNA and 7 ug of mRNA. A directionally cloned cDNA library in the pCMVSPORT6 vector was constructed at Life Technologies (Rockville, MD; now part of Invitrogen Corp), essentially following the protocols of the SuperScript Plasmid System (Invitrogen Corp). <http://www.invitrogen.com/>". The library code designation was cs. For this library, cDNA inserts were cloned into the NotI/MluI sites of the vector. EST analysis was performed on the unamplified library at the NIH Intramural Sequencing Center (NISC)."
FEATURES	
SOURCE	
ORIGIN	
Query Match	28.6%; Score 638; DB 6; Length 639;
Best Local Similarity	99.8%; Pred. No. 3.6e-150;
Matches 638; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
QY	12 ACCAGCTAGTGGCCAGACTTTCTGTGGATCATCGAGCCACCTGMAACCCACCTGAC 71
DB	1 ACCAGCTAGTGGCCAGACTTTCTGTGGATCATCGAGCCACCTGMAACCCACCTGAC 60
QY	72 CCAAGCCACCTAGTGGCCAGCCACCTGTCGATGACCACTTACATCAACAAGCAAGTG 131
DB	61 CCAAGCCACCTAGTGGCCAGCCACCTGTCGATGACCACTTACATCAACAAGCAAGTG 120
QY	132 GCTATAGCCGCGCTTAGGCTCTTCGCCGCTGCTGTCGTGCGCGGAGAGCATCTAC 191
DB	121 GCTATAGCCGCGCTTAGGCTCTTCGCCGCTGCTGTCGTGCGGAGAGCATCTAC 180

Accession	Version	Source	Organism	Reference	Authors	Title	Journal	Medline	Pubmed	Comment
QY	192	AAAGTGCATATGGACGAGTTCTTAAATCTTCCTGCTGCTACATCAATCCGCTTATT	251							
Db	181	AAAGTGCATATGGACGAGTTCTTAAATCTTCCTGCTGCTACATCAATCCGCTTATT	240							
QY	252	TATAGGCTGGCCCTCAACGGAAGAACAAACAGCTGATGTTTGAGAACTGACTCTGTAATGC	311							
Db	241	TATAGGCTGGCCCTCAACGGAAGAACAAACAGCTGATGTTTGAGAACTGACTCTGTAATGC	300							
QY	312	GACAGCTACATCCAGCTCATATCCCATTTCTTCTGCTGGAGCTTCTACGTACAGCTGTGC	371							
Db	301	GACAGCTACATCCAGCTCATATCCCATTTCTTCTGCTGGAGCTTCTACGTACAGCTGTGC	360							
QY	372	GTGACCCGCTGGTGGTGAACCAAGTACAGAACCTGCGCTGGCCCGACCGGCTCATGAGCTG	431							
Db	361	GTGACCCGCTGGTGGTGAACCAAGTACAGAACCTGCGCTGGCCCGACCGGCTCATGAGCTG	420							
QY	432	GTCGTGGGCTTTCGTGCAAGGCAAGAACGAGCAAGCGCGCTGTCGCGACGCTCATC	491							
Db	421	GTCGTGGGCTTTCGTGCAAGGCAAGAACGAGCAAGCGCGCTGTCGCGACGCTCATC	480							
QY	492	CGTACGCGCAACTGTGGCAACGTGCTCATCTGCGACGCTCAGACCGCAGTTCACAG	551							
Db	481	CGTACGCGCAACTGTGGCAACGTGCTCATCTGCGACGCTCAGACCGCAGTTCACAG	540							
QY	552	CGCTTCCCGACGCGCCAGCACTGTGTGCAAGCAGAGCGTTTAACTCCGCGGAGAACACAG	611							
Db	541	CGCTTCCCGACGCGCCAGCACTGTGTGCAAGCAGAGCGTTTAACTCCGCGGAGAACACAG	600							
QY	612	CAGTTGGAGAAACTGAGCTTACACACACAACTGTTCTGG	650							
Db	601	CAGTTGGAGAAACTGAGCTTACACACACAACTGTTCTGG	639							
RESULT 11										
LOCUS	CA397981									
DEFINITION	CA397981	636 bp	mRNA	linear	EST 06-NOV-2002					
	cg988c05.y1		Human Retinal pigment epithelium/choroid cDNA							
	(un-normalized, unamplified):	cg	Homo sapiens	cdna	clone	cg988c05				
	5', mRNA sequence.									
VERSION	CA397981									
KEYWORDS	CA397981.1	GI:24735789								
SOURCE	EST.									
ORGANISM	Homo sapiens	(human)								
	Homo sapiens									
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
	1 (bases 1 to 636)									
	Wiatow G., Bernstein S.L., Wyatt M.K., Farris R.N., Behal A.,									
	Touchman J.W., Bouffard G., Smith D. and Peterson K.									
	Expressed sequence tag analysis of human RPE/choroid for the									
	NEI/Bank Project: Over 6000 non-redundant transcripts, novel genes									
	and splice variants									
	Mol. Vis. 8 (4), 205-220 (2002)									
JOURNAL	22103460									
MEDLINE	12107410									
PUBMED	Contact: Wiatow G									
COMMENT	Section on Molecular Structure and Function									
	National Eye Institute									
	6/331, NIH, Bethesda, MD 20892-2740, USA									
	Tel: 301 402 3452									
	Fax: 301 496 0078									
	Email: grame@helix.nih.gov									
	Plate: 98 row: C column: 05									
	Seq primer: M13RP1 reverse primer (ABI).	</								

/clone lib="Human Retinal pigment epithelium/choroid cDNA (un-normalized, unamplified): cs"
 /note="Organ: Eye; Vector: pCMVSPORT6; Two different donor eyes (75-80 years old) yielded approximately 600 mg of dissected RPE/choroid tissue. This in turn yielded 340 ug of total RNA and 7 ug of mRNA. A directionally cloned cDNA library in the pCMVSPORT6 vector was constructed at Life Technologies (Rockville, MD; now part of Invitrogen Corp), essentially following the protocols of the Superscript Plasmid System (Invitrogen Corp).
 <http://www.invitrogen.com/>. The library code designation was cs. For this library, cDNA inserts were cloned into the NotI/MluI sites of the vector. EST analysis was performed on the unamplified library at the NIH Intramural Sequencing Center (NISC)."

ORIGIN

```

Query Match      28.3%; Score 631.2; DB 6; Length 636;
Best Local Similarity 99.5%; Pred. No. 1.9e-148;
Matches 633; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1125 CCCGAGCCACAGCCCCCTACACAGCTGCTCCGCCCATGTTCCGTCGAGCCCTTTATG 1184
DB 1 CCCGAGCCACAGCCCCCTACACAGCTGCTCCGCCCATGTTCCGTCGAGCCCTTTATG 60
QY 1185 GCGTCCACCTTCAACATCAGCTGGAACAAAGAGATGAGTTCCAGCCCAATCAGAG 1244
DB 61 GCGTCCACCTTCAACATCAGCTGGAACAAAGAGATGAGTTCCAGCCCAATCAGAG 120
QY 1245 GACGAGGAGATGCTCAGGCTGGCATCATGTCGCGCCCTTCCATGAGCTTCAGTCCCATGAT 1304
DB 121 GACGAGGAGATGCTCAGGCTGGCATCATGTCGCGCCCTTCCATGAGCTTCAGTCCCATGAT 180
QY 1305 CACCATCTCTCCAGGGCAAACTCAAGAGACAAACTACTGTGAGCCCAAGAGGAATCCCTT 1364
DB 181 CACCATCTCTCCAGGGCAAACTCAAGAGACAAACTACTGTGAGCCCAAGAGGAATCCCTT 240
QY 1365 CTCGACGAGGGCGTGGCCCAAAACCAAGAGCAGCCCAAGACGATTAAGGGGCGCAGGAA 1424
DB 241 CTCGACGAGGGCGTGGCCCAAAACCAAGAGCAGCCCAAGACGATTAAGGGGCGCAGGAA 300
QY 1425 GACAAACAAGGCGCTGGAAGCTTAAGGCTGTGAGCGCTTCAAGTGTGGCCCACTGTATCAG 1484
DB 301 GACAAACAAGGCGCTGGAAGCTTAAGGCTGTGAGCGCTTCAAGTGTGGCCCACTGTATCAG 360
QY 1485 AGGCGCAGGCTACTACAGTGGCCCAACAAGCGCCCTCAGCCCACTCCCATGTTCTTCCC 1544
DB 361 AGGCGCAGGCTACTACAGTGGCCCAACAAGCGCCCTCAGCCCACTCCCATGTTCTTCCC 420
QY 1545 CTGAACCATCAGGCGCGTCAAGAGCTTCAAGTGTCAAGGCAATGACACCAAGACAAA 1604
DB 421 CTGAACCATCAGGCGCGTCAAGAGCTTCAAGTGTCAAGGCAATGACACCAAGACAAA 480
QY 1605 AGCTTAAAGACTGTGAGTTCTGGGGCCAAAGAAAGTTTGAATTTGCTCTCAGAGAGGAT 1664
DB 481 AGCTTAAAGACTGTGAGTTCTGGGGCCAAAGAAAGTTTGAATTTGCTCTCAGAGAGGAT 540
QY 1665 GGGGCGCTTGAATGAGGACCAAGAGTATCTCAAGTGAAGGAGAAAGCTGTGAAGTTAAC 1724
DB 541 GGGGCGCTTGAATGAGGACCAAGAGTATCTCAAGTGAAGGAGAAAGCTGTGAAGTTAAC 600
QY 1725 CTGACGATATGCGACAGATCCCGGAAATCACCCTC 1760
DB 601 CTGACGATATGCGACAGATCCCGGAAATCACCCTC 636

```

RESULT 12

BU731149/c 651 bp mRNA linear EST 09-OCT-2002
 LOCUS BU731149
 DEFINITION UI-E-CII-afq-j-20-0-UI-s1 UI-E-CII Homo sapiens cDNA clone
 ACCESSION BU731149
 VERSION BU731149.1 GI:23655752

KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 651)
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 PUBMED 8889548
 COMMENT Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@iowa.edu
 Tissue Procurement: Dr. Gregg Hageman
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).
 Seq primer: M13 FORWARD
 POLYA=Yes.

FEATURES

source
 Location/Qualifiers
 1..651
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-E-CII-afq-j-20-0-UI"
 /tissue_type="RPE and Choroid"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /clone_lib="UI-E-CII"
 /note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-E-CII is a normalized cDNA library containing the following tissue(s): RPE and Choroid. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is ACTTA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI).
 TAG_TISSUE=RPE and Choroid
 TAG_LIB=UI-E-CII
 TAG_SEQ=ACCTA"

ORIGIN

```

Query Match      28.1%; Score 626.6; DB 5; Length 651;
Best Local Similarity 99.2%; Pred. No. 2.8e-147;
Matches 629; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1390 ACAAGGACGCAACAGAACTTAGGGCCAGAGACAAACAAGGCTGGAAGCTTAAG 1449
DB 651 ACAAGGACGCAACAGAACTTAGGGCCAGAGACAAACAAGGCTGGAAGCTTAAG 592
QY 1450 CTGTGAGGCTTTAAAGCTTGGGCCCACTGTATCAGAGGCCAGGCTACTAAGTCCAC 1509
DB 591 CTGTGAGGCTTTAAAGCTTGGGCCCACTGTATCAGAGGCCAGGCTACTAAGTCCAC 532
QY 1510 AGAGCGCCCTCAGGCCCACTCCCATGTTCTTCCCGCTAGAACCATCAGCGCCGCTAAGC 1569
DB 531 AGAGCGCCCTCAGGCCCACTCCCATGTTCTTCCCGCTAGAACCATCAGCGCCGCTAAGC 472

```

QY 1570 TTCACAGTGTACAGGATAGACACCAAGAAAGCTTAAGACTGTGAGTTCTGGGG 1629
| | | | |
Db 471 TTCACAGTGTACAGGATAGACACCAAGAAAGCTTAAGACTGTGAGTTCTGGGG 412
| | | | |
QY 1630 CCAAGAAAAGTTTGAATGCTCTCAGAGAGCGATGGGGCTTGAATGAGCACCAGAG 1689
| | | | |
Db 411 CCAAGAAAAGTTTGAATGCTCTCAGAGAGCGATGGGGCTTGAATGAGCACCAGAG 352
| | | | |
QY 1690 TATCTCAAGTGAAGAGAAAAGCTGTGAGTTAACTGACCGATATGCGACAGATCCCGG 1749
| | | | |
Db 351 TATCTCAAGTGAAGAGAAAAGCTGTGAGTTAACTGACCGATATGCGACAGATCCCGG 292
| | | | |
QY 1750 AAAATCACTCAAGAAACCTTTGGAAACATACCAACCAATATACATCACTCACTCAAG 1809
| | | | |
Db 291 AAAATCACTCAAGAAACCTTTGGAAACATACCAACCAATATACATCACTCAAG 232
| | | | |
QY 1810 ATCAATGATGATCTTATTTGGGCTTGGAAAACAGGGATGAAGACATTTCTTACCTGCTT 1869
| | | | |
Db 231 ATCAATGATGATCTTATTTGGGCTTGGAAAACAGGGATGAAGACATTTCTTACCTGCTT 172
| | | | |
QY 1870 CCTAATGGGATGCTTGGCCAGCCAGTCTCTACCTGTGTGTACACCGACGAGACACTGA 1929
| | | | |
Db 171 CCTAATGGGATGCTTGGCCAGCCAGTCTCTACCTGTGTGTACACCGACGAGACACTGA 112
| | | | |
QY 1930 TCCAGTCAAGCCATACAGCTGTCCACATGAAAGCTGTCTTACCAACAGCCTGAATCA 1989
| | | | |
Db 111 TCCAGTCAAGCCATACAGCTGTCCACATGAAAGCTGTCTTACCAACAGCCTGAATCA 52
| | | | |
QY 1990 AATGTTAGCTTAATAGATAAAAAATCCCACTA 2023
| | | | |
Db 51 AATGTTAGCTTAATAGATAAAAAATCCCACTA 18
| | | | |

RESULT 13
BM663028/c
LOCUS
DEFINITION
U-E-C10-aad-h-10-0-UI.81 UI-E-C10 Homo sapiens cDNA clone
BM663028
U-E-C10-aad-h-10-0-UI 3', mRNA sequence.
ACCSSION
BM663028.1 GI:18968017
VERSION
EST.
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1 (bases 1 to 592)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
JOURNAL
MEDLINE
PUBMED
8889548
COMMENT
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 Forward
POLYA=yes.

FEATURES
Source
Location/Qualifiers
1..592
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

/clone="UI-E-C10-aad-h-10-0-UI"
/issue_type="RPE and Choroid"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_id="UI-E-C10"
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site 1: BcoR I; Site 2: Not I;
UI-E-C10 is a cDNA library containing the following
tissue(s): RPE and Choroid. The library was constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an BcoR I adaptor, digested
with Not I, and cloned directionally into pT73-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (drr)18 tail. The
sequence tag for this library is ACCTA. This library was
created for the program, Gene Discovery in the Visual
System, supported by National Eye Institute (NEI).
TAG_LIB=UI-E-C10
TAG_SEQ=ACCTA"

ORIGIN
Query Match 26.1%; Score 582.4; DB 4; Length 592;
Best Local Similarity 99.0%; Pred.No. 4.3e-116;
Matches 586; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1627 GGGCCAAAGAAAAGTTTGAATGCTCTCAGAGAGCGATGGGGCTTGAATGAGCACCAG 1686
| | | | |
Db 592 GGGCCAAAGAAAAGTTTGAATGCTCTCAGAGAGCGATGGGGCTTGAATGAGCACCAG 533
| | | | |
QY 1687 AAGTATCTCAAGTGAAGAGAAAAGCTGTGAGTTAACTGACCGATATGCGACAGATCC 1746
| | | | |
Db 532 AAGTATCTCAAGTGAAGAGAAAAGCTGTGAGTTAACTGACCGATATGCGACAGATCC 473
| | | | |
QY 1747 CCGAAAATCACTCAAGAAACCTTTGGAAACATACCAACCAATATACATCACTCACTCA 1806
| | | | |
Db 472 CCGAAAATCACTCAAGAAACCTTTGGAAACATACCAACCAATATACATCACTCACTCA 413
| | | | |
QY 1807 AAGATCAATGATCTTATTTGGGCTTGGAAAACAGGGATGAAGACATTTCTTACCTG 1866
| | | | |
Db 412 AAGATCAATGATCTTATTTGGGCTTGGAAAACAGGGATGAAGACATTTCTTACCTG 353
| | | | |
QY 1867 CTTCCTAATGGGATGCTTGGCCAGCCAGTCTCTACCTGTGTGTACACCGACGAGAC 1926
| | | | |
Db 352 CTTCCTAATGGGATGCTTGGCCAGCCAGTCTCTACCTGTGTGTACACCGACGAGAC 293
| | | | |
QY 1927 TGATCCAGTCAAGCCATACAGCTGTCCACATGGAAGAGTGTCTTACCAACAGCCTGAA 1986
| | | | |
Db 292 TGATCCAGTCAAGCCATACAGCTGTCCACATGGAAGAGTGTCTTACCAACAGCCTGAA 233
| | | | |
QY 1987 TCAATGTTAGCTTAATAGATAAAAAATCCAGACTACTTCAGCCTTAAATGCTTTTAT 2046
| | | | |
Db 232 TCAATGTTAGCTTAATAGATAAAAAATCCAGACTACTTCAGCCTTAAATGCTTTTAT 173
| | | | |
QY 2047 TCATAAAAGCTGGAAGCTAGACTGAACCATTTGAAACATTTAAGTCAAGCTGGATT 2106
| | | | |
Db 172 TCATAAAAGCTGGAAGCTAGACTGAACCATTTGAAACATTTAAGTCAAGCTGGATT 113
| | | | |
QY 2107 CAGAGTGGGAAACCTTGAATGCTATCTGAATCCAAAGACAGCAACCTTGTATATGCGC 2166
| | | | |
Db 112 CAGAGTGGGAAACCTTGAATGCTATCTGAATCCAAAGACAGCAACCTTGTATATGCGC 53
| | | | |
QY 2167 CAATCTAATGATTTAATTAATCAATATCTCGTTAAAAAAGAAAAA 2218
| | | | |
Db 52 CAATCTAATGATTTAATTAATCAATATCTCGTTAAAAAAGAAAAA 1
| | | | |

RESULT 14
BE410951
LOCUS
BE410951 773 bp mRNA linear EST 21-JUL-2000

DEFINITION 601303662F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3638175 5', mRNA sequence.
 BE410951
 VERSION BE410951.1 GI:9347401
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 773)
 NIH-MGC http://mgc.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cs9abs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINTL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINTL at: image.lnl.gov
 Plate: L1CM38 row: 1 column: 16
 High quality sequence stop: 662.
 Location/Qualifiers
 1..773
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3638175"
 /tissue_type="choriocarcinoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: Placenta; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

FEATURES

source

ORIGIN

Query Match 25.9%; Score 577; DB 2; Length 773;
 Best Local Similarity 99.0%; Pred. No. 1,1e-134;
 Matches 591; Conservative 0; Mismatches 5; Indels 1; Gaps 1;
 1253 GGATGCTACCGCTGCGATCATTTGGCGCTTCTTCTAGGCTGATGCCATGATCCATCC 1312
 1 GGATGCTACCGCTGCGATCATTTGGCGCTTCTTCTAGGCTGATGCCATGATCCATCC 60
 1313 TCCCAAGGCAACTCAAGACCAACTACTGTGCGCCCAAGAGGGAATCCCTTCCACCA 1372
 61 TCCCAAGGCAACTCAAGACCAACTACTGTGCGCCCAAGAGGGAATCCCTTCCACCA 120
 1373 GGGCGCTGCC-AAAACCAAGAGGAGCAAGCAAGGAGGAGGAGGAGGAGGAGGAGGAG 1431
 121 GGGCGCTGCCAAAACCAAGAGGAGGAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180
 1432 AGGCTTGGAAGCTTAAGAGCTGTGAGCGCTTCAAGTGTGCGCCCACTGATCAGAGGCCG 1491
 181 AGGCTTGGAAGCTTAAGAGCTGTGAGCGCTTCAAGTGTGCGCCCACTGATCAGAGGCCG 240
 1492 GCTACTACATGCCCCCAAGAGCGCCCTTCAAGTGTGCGCCCACTGATCAGAGGCCG 1551
 241 GCTACTACATGCCCCCAAGAGCGCCCTTCAAGTGTGCGCCCACTGATCAGAGGCCG 300
 1552 CATCAGCGCGCTCAAGAGCTTCAAGTGTGAGCGGCAAGGAGGAGGAGGAGGAGGAGGAG 1611
 301 CATCAGCGCGCTCAAGAGCTTCAAGTGTGAGCGGCAAGGAGGAGGAGGAGGAGGAGGAG 360
 1612 AGACTGTGAGTTTGGGGCCCAAGAAAGTTTGAATTTGCTTCAGAGAGGAGTGGGGCT 1671
 361 AGACTGTGAGTTTGGGGCCCAAGAAAGTTTGAATTTGCTTCAGAGAGGAGTGGGGCT 420

QY 1672 TGATGAGACCCAGAAATATCTCAAGTGAAGAGAAAAGTGTGAATTAACCTGACGG 1731
 DB 421 TGATGAGACCCAGAAATATCTCAAGTGAAGAGAAAAGTGTGAATTAACCTGACGG 480
 QY 1732 ATATGCGAGATCCCGGAAATATCACTCAAGAAAGGAGGAGGAGGAGGAGGAGGAGGAG 1791
 DB 481 ATATGCGAGATCCCGGAAATATCACTCAAGAAAGGAGGAGGAGGAGGAGGAGGAGGAG 540
 QY 1792 TACACATCACTCAAGAGTCAATGATGATGATCTTATTTGGGCTTGAAGAAACAGGATG 1848
 DB 541 TACACATCACTCAAGAGTCAATGATGATGATCTTATTTGGGCTTGAAGAAACAGGATG 597

RESULT 15

BM707948

LOCUS

DEFINITION

BM707948

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

FEATURES

source

1..585
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-E-C11-aft-g-11-0-UI"
 /tissue_type="RPE and Choroid"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-E-C11 is a normalized cDNA library containing the constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. The library was synthesized with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag (dT)18 tail. The sequence tag for this library is ACCTA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

ORIGIN

Query Match 25 8%; Score 575; DB 4; Length 585;

Best Local Similarity 98.8%; Pred. No. 3.2e-134;

Matches 578; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
QY 1556 AGCCCGCTCAAAAGCTTCACAGTGTACAGCATAGACACCAAGACAAAAGCTTAAAGAC 1615
    |||
Db 1 AGCCCGCTCAAAAGCTTCACAGTGTACAGCATAGACACCAAGACAAAAGCTTAAAGAC 60
QY 1616 TGTAGTTCGTGGGCGCAAGAAAGTTTGAATGCTCTCAGAGAGCGATGCGGCTTGAT 1675
    |||
Db 61 TGTAGTTCGTGGGCGCAAGAAAGTTTGAATGCTCTCAGAGAGCGATGCGGCTTGAT 120
QY 1676 GGAGACCCCAAGATATCTCAAGTGAAGAGAAAAGTGTGAAGTTTAACTGACGATAT 1735
    |||
Db 121 GGAGACCCCAAGATATCTCAAGTGAAGAGAAAAGTGTGAAGTTTAACTGACGATAT 180
QY 1736 GCCAGAGTCCCGCAAAATCACTCAAGAAACCTTTGGACAATCACCAACCAACATACA 1795
    |||
Db 181 GCCAGAGTCCCGCAAAATCACTCAAGAAACCTTTGGACAATCACCAACCAACATACA 240
QY 1796 CACTACACTCAAGATCATGATGCTTATTTGGGCTTTGAAAAACAGGATGAAGACA 1855
    |||
Db 241 CACTACACTCAAGATCATGATGCTTATTTGGGCTTTGAAAAACAGGATGAAGACA 300
QY 1856 TTCTTAACCTGCTTCTTAATGGGATGCTTGGCCAGCCAGGTCTCACTGTGTATAC 1915
    |||
Db 301 TTCTTAACCTGCTTCTTAATGGGATGCTTGGCCAGCCAGGTCTCACTGTGTATAC 360
QY 1916 CAGCAGGCACTGATCCAGTCAAGCCATACAGCTGTCCACACTGAAGAACGTGCTTAC 1975
    |||
Db 361 CAGCAGGCACTGATCCAGTCAAGCCATACAGCTGTCCACACTGAAGAACATGCTTAC 420
QY 1976 AACAGCCTGAATCAATGTTAGCTTAATAGATAAAATCCAGACTACTTCAGCCTTTA 2035
    |||
Db 421 AACAGCCTGAATCAATGTTAGCTTAATAGATAAAATCCAGACTACTTCAGCCTTTA 480
QY 2036 ATGCTTTTATTCATTAATAAACTGTGAAGCTAGACTGAACATTTGAAAACATTTA 2095
    |||
Db 481 ATGCTTTTATTCATTAATAAACTGTGAAGCTAGACTGAACATTTGAAAACATTTA 540
QY 2096 GACTCTGATTCAGAGTCGGGACCCCTTAGTTCTATCTGAATCCA 2140
    |||
Db 541 GACTCTGATTCAGAGTCGGGACCCCTTAGTTCTATCTGAATCCA 585
```

Search completed: March 26, 2005, 13:32:03
Job time : 7175.13 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2005, 21:57:50 ; Search time 1155.53 seconds
(without alignments)
11419.125 Million cell updates/sec

Title: US-09-622-964A-2

Perfect score: 2229

Sequence: 1 caggagagcaccacagccta.....aaaaaaaaaaaaaaaaaaaa 2229

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2229	100.0	2229	2 AA221227	AA221227 Human CG1
2	2013	90.3	2429	2 AA221228	AA221228 Human CG1
3	1984.8	89.0	2404	12 AD084435	Adq84435 Human tum
4	1984.8	89.0	2404	13 AD083262	Adq83262 Human tum
5	1754.8	78.7	1758	8 AB280972	Ab280972 Human des
6	1247.2	56.0	1263	2 AA99722	AA99722 Human adu
7	1011.2	45.4	1238	9 ADA44960	Ada44960 Human pol
8	871.2	39.1	1916	2 AA221229	AA221229 Mouse CG1
9	642.2	28.8	16125	2 AA221226	AA221226 Human CG1
10	642.2	28.8	16550	5 AB14559	Ab14559 Human ner
11	642.2	28.8	18530	5 AB14557	Ab14557 Human ner
12	640.6	28.7	18537	5 AB14558	Ab14558 Human ner
13	587.8	26.4	1717	6 AB189697	Ab189697 Human pol
14	577.2	25.9	2112	6 AB061177	Ab061177 FLJ20132
15	577.2	25.9	2137	8 AB280973	Ab280973 Human des
16	576.8	25.9	1530	8 AB280997	Ab280997 Human des
17	476.4	21.4	7108	5 AB14556	Ab14556 Human ner
18	459.2	20.6	2500	11 ADM02614	Adm02614 Human CDN
19	448.2	20.1	1948	9 ACC59900	ACC59900 Human REM
20	446.2	20.0	1422	8 AB280975	Ab280975 Human des

21	423	19.0	2028	8 AB280974	Ab280974 Human des
22	350.4	15.7	1326	8 ACA03889	ACA03889 CDNA upre
23	314.8	14.1	2861	4 AB110793	Ab110793 Drosophill
24	294.6	13.2	305	5 AB11094	Ab11094 Human ner
25	275.8	12.4	1526	5 AAS72796	Aas72796 DNA encod
26	252	11.3	1608	4 AB112609	Ab112609 Drosophill
27	243.4	10.9	630	13 ADQ53985	Adq53985 Novel can
28	234.8	10.5	424	2 AA986976	AA986976 EST clone
29	219.2	9.8	1592	10 ADC29982	Adc29982 Human nov
30	218.6	9.8	10760	4 AB110792	Ab110792 Drosophill
31	215.4	9.7	1345	4 AB112637	Ab112637 Drosophill
32	215.4	9.7	1350	4 AAF76848	Aaf76848 Human sec
33	213	9.6	620	5 AD145614	Ad145614 Human ova
34	199	8.9	2196	4 AB116195	Ab116195 Drosophill
35	189.2	8.5	1292	11 ACN89920	Acn89920 Breast ca
36	179.4	8.0	3592	4 AB112636	Ab112636 Drosophill
37	175	7.9	1119	8 AB271972	Ab271972 Human fer
38	175	7.9	1198	6 ABK83853	Abk83853 Human CDN
39	175	7.9	1198	6 ABN95676	Abn95676 Gene #217
40	175	7.9	1234	5 AAS67682	Aas67682 DNA encod
41	175	7.9	1279	5 AAS85665	Aas85665 DNA encod
42	175	7.9	2183	5 AAS91587	Aas91587 DNA encod
43	139.8	6.3	3785	4 AB112608	Ab112608 Drosophill
44	135.2	6.1	539	12 ACH74388	Ach74388 Human gen
45	134.2	6.0	235	12 ACH88088	Ach88088 Human gen

ALIGNMENTS

RESULT 1	
AA221227	
ID	AA221227 standard; CDNA; 2229 BP.
XX	
XX	AA221227;
AC	
XX	
DT	22-NOV-1999 (first entry)
XX	
DE	Human CGICE short form CDNA sequence.
XX	
KW	CGICE, Beas' s macular dystrophy; mutation; diagnosis; detection; BMD;
KM	age-related macular dystrophy; ss.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	105..1862
FT	/*tag= a
FT	/product= "CGICE long form protein"
FT	/transl_except= (pos:465..467,aa:Ser)
XX	
PN	W09943695-A1.
XX	
PD	02-SEP-1999.
XX	
PF	22-FEB-1999; 99WO-US003790.
XX	
PR	25-FEB-1998; 98US-0075941P.
XX	
PR	18-DEC-1998; 98US-0112926P.
XX	
PA	(MERI) MERCK & CO INC.
XX	
XX	(UYUP-) UNIV UPPSALA.
PI	Petrukhin K, Caskey CT, Metzker M, Madelius C;
XX	
DR	WPI, 1999-540560/45.
XX	
DR	P-PSDB; AAY29953.
XX	
PT	Human and mouse polynucleotides encoding CGICE polypeptides.
XX	
PS	Claim 2; Fig 2; 67bp; English.
XX	
CC	The present sequence represents the human CGICE CDNA sequence, which when

CC mutated is responsible for Best's macular dystrophy (BMD) .
CC Polynucleotides encoding CG1CE are useful for diagnosing whether a
CC patient carries a mutation in the CG1CE gene. Normal and mutated CG1CE
CC proteins are useful for identifying activators and/or inhibitors of these
CC proteins, in order to treat BMD. The CG1CE gene offers a simpler and
CC cheaper method of diagnosing BMD without the need for the presence of the
CC patient. The gene may also be useful to discovering the genetic cause of
CC age-related macular dystrophy

SQ Sequence 2229 BP; 575 A; 646 C; 532 G; 476 T; 0 U; 0 Other;

Query Match	100.0%	Score 2229;	DB 2;	Length 2229;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2229; Conservative	0;	Mismatches	0;	Gaps 0

QY	1	CAGGAGTCCCAACAGCTTATGTGCGCAGACCTTCTGTGGATCATCGGACCCACTTGAA	60
Db	1	CAGGAGTCCCAACAGCTTATGTGCGCAGACCTTCTGTGGATCATCGGACCCACTTGAA	60
QY	61	CCCCACTTGAAGCCCAAGGCCACTGTGCGAGGCCCACTGTGCGCATGACATCATCTTACA	120
Db	61	CCCCACTTGAAGCCCAAGGCCACTGTGCGAGGCCCACTGTGCGCATGACATCATCTTACA	120
QY	121	CAAGCCAAGTGCTAAATGCGCGTTAGGCTCTCTTCCCGCTGCGCTGTGTCGGCGG	180
Db	121	CAAGCCAAGTGCTAAATGCGCGTTAGGCTCTCTTCCCGCTGCGCTGTGTCGGCGG	180
QY	181	GGAGCATCTACAAAGCTGTATATGTCGAGTTCTTAATCTTCTGTCTGTCTATCATCA	240
Db	181	GGAGCATCTACAAAGCTGTATATGTCGAGTTCTTAATCTTCTGTCTGTCTATCATCA	240
QY	241	TCCGCTTATTTATAGGCTGGCCCTCAGCGGAAAGAACAAAGCTGATGTTTGAAACTGA	300
Db	241	TCCGCTTATTTATAGGCTGGCCCTCAGCGGAAAGAACAAAGCTGATGTTTGAAACTGA	300
QY	301	CTCTGTATTTGGACAGCTTACATCCAGCTCATCCCATTTCTCTGTGCTGGACTTCTACG	360
Db	301	CTCTGTATTTGGACAGCTTACATCCAGCTCATCCCATTTCTCTGTGCTGGACTTCTACG	360
QY	361	TGACCGTGTGCTGACCCGCTGTGTGGAACACAGTACGAGAACTTGCGCCGACCCGC	420
Db	361	TGACCGTGTGCTGACCCGCTGTGTGGAACACAGTACGAGAACTTGCGCCGACCCGC	420
QY	421	TCATGAGCTGTGTGCTGCGGCTTCTGTGAAAGCAAGAAACAGCAAGCGCGCTGCGGGC	480
Db	421	TCATGAGCTGTGTGCTGCGGCTTCTGTGAAAGCAAGAAACAGCAAGCGCGCTGCGGGC	480
QY	481	GCACGCTCATCGCTACCGCAACCTTGCGCAACGTCTCATCTCTGCGCAGCGTCAGCACG	540
Db	481	GCACGCTCATCGCTACCGCAACCTTGCGCAACGTCTCATCTCTGCGCAGCGTCAGCACG	540
QY	541	CAGTCTACAGCGCTTCCCGACGCGCCAGACCTGTGTCAAGCAGAGCTTATGACTCCGG	600
Db	541	CAGTCTACAGCGCTTCCCGACGCGCCAGACCTGTGTCAAGCAGAGCTTATGACTCCGG	600
QY	601	CAGAAACAAGCAAGTTGGAGAACTGACCTTACACAACAACATGTTCTGTGCTGCGCGG	660
Db	601	CAGAAACAAGCAAGTTGGAGAACTGACCTTACACAACAACATGTTCTGTGCTGCGCGG	660
QY	661	TGTGTTTGGCCAACTGTCAATGAAGCGTGTGCTTGAAGTTCGAATCCGGGACCTTATCC	720
Db	661	TGTGTTTGGCCAACTGTCAATGAAGCGTGTGCTTGAAGTTCGAATCCGGGACCTTATCC	720
QY	721	TGCTTCAGAGCTGCTGAAACGAGATGGAACACTTGCGTATCACTCAATGTGACACCTTGATG	780
Db	721	TGCTTCAGAGCTGCTGAAACGAGATGGAACACTTGCGTATCACTCAATGTGACACCTTGATG	780
QY	781	CTTACGACTGATTAAGTATCCCACTGTGTATACAGAGTGTGATCTGTGGCGGTGACAA	840
Db	781	CTTACGACTGATTAAGTATCCCACTGTGTATACAGAGTGTGATCTGTGGCGGTGACAA	840
QY	841	GCCTCTTCTGACCTTGTCTAGTTGGGCGGCAAGTTTCTGAACCAAGCCAAGCCTTACCTTG	900
Db	841	GCCTCTTCTGACCTTGTCTAGTTGGGCGGCAAGTTTCTGAACCAAGCCAAGCCTTACCTTG	900

Db	841	GCCTCTTCTGACTTGTGCTAGTGGGCGCAGTTTCTGAACCCAGCCAAAGGCGTACCTTG	900													
Qy	901	GCATAGAGCTGGA	CCTGATTGTGCCCGCTTTCACGTTCCGACAGTTCCTTCTATAGTTG	960												
Db	901	GCATAGAGCTGGA	CCTGATTGTGCCCGCTTTCACGTTCCGACAGTTCCTTCTATAGTTG	960												
Qy	961	GTGGGCTGAAGGTGGCAGAGACACTATCAACCCCTTTGGAGAGATGATGATTTTG	1021	961	GTGGGCTGAAGGTGGCAGAGACACTATCAACCCCTTTGGAGAGATGATGATTTTG	1021										
Db	1021	AGAACA	CTGATTTGTGACAGAGAA	TTTGACAGGTG	CCCTGTGGCTGGAGATGAGATGC	1081	1021	AGAACA	CTGATTTGTGACAGAGAA	TTTGACAGGTG	CCCTGTGGCTGGAGATGAGATGC	1081				
Qy	1021	AGAACA	CTGATTTGTGACAGAGAA	TTTGACAGGTG	CCCTGTGGCTGGAGATGAGATGC	1081	1021	AGAACA	CTGATTTGTGACAGAGAA	TTTGACAGGTG	CCCTGTGGCTGGAGATGAGATGC	1081				
Db	1081	ACCAAGAC	CTGCTCGGATGGAGCCGGA	CATGTA	CTGAAATTAAGCCCGACCAAGCCCC	1141	1081	ACCAAGAC	CTGCTCGGATGGAGCCGGA	CATGTA	CTGAAATTAAGCCCGACCAAGCCCC	1141				
Qy	1141	CTTACA	CAGCTGCTTCCGCGCAAGTTCG	CGTAGCCCTTTATTTGGCTTCA	CCTTCAACA	1201	1141	CTTACA	CAGCTGCTTCCGCGCAAGTTCG	CGTAGCCCTTTATTTGGCTTCA	CCTTCAACA	1201				
Db	1141	CTTACA	CAGCTGCTTCCGCGCAAGTTCG	CGTAGCCCTTTATTTGGCTTCA	CCTTCAACA	1201	1141	CTTACA	CAGCTGCTTCCGCGCAAGTTCG	CGTAGCCCTTTATTTGGCTTCA	CCTTCAACA	1201				
Qy	1201	TCAGCCTGAA	CAAAAGAGAGATGAGTTCCAGCCCAATCAGAGAGACGAGAGATGCTC	1261	1201	TCAGCCTGAA	CAAAAGAGAGATGAGTTCCAGCCCAATCAGAGAGACGAGAGATGCTC	1261								
Db	1261	AGCGTGGCAT	CAATTTGGCGGCTTCC	TAGGGCTTCACTCCATGATCA	CCATCTCCAGGG	1321	1261	AGCGTGGCAT	CAATTTGGCGGCTTCC	TAGGGCTTCACTCCATGATCA	CCATCTCCAGGG	1321				
Qy	1321	CAAACTCA	AGAGACCAAACTACTGTGTGGCCCAAGAGGGAATCCCTTCTCACAGAGGCGCTGC	1381	1321	CAAACTCA	AGAGACCAAACTACTGTGTGGCCCAAGAGGGAATCCCTTCTCACAGAGGCGCTGC	1381								
Db	1321	CAAACTCA	AGAGACCAAACTACTGTGTGGCCCAAGAGGGAATCCCTTCTCACAGAGGCGCTGC	1381	1321	CAAACTCA	AGAGACCAAACTACTGTGTGGCCCAAGAGGGAATCCCTTCTCACAGAGGCGCTGC	1381								
Qy	1381	CCAAAA	CCACAAGCAGCCAAACA	GAACGTTAAGGGGCGCAGAGAACAAACAAGGCTTGA	1441	1381	CCAAAA	CCACAAGCAGCCAAACA	GAACGTTAAGGGGCGCAGAGAACAAACAAGGCTTGA	1441						
Db	1381	CCAAAA	CCACAAGCAGCCAAACA	GAACGTTAAGGGGCGCAGAGAACAAACAAGGCTTGA	1441	1381	CCAAAA	CCACAAGCAGCCAAACA	GAACGTTAAGGGGCGCAGAGAACAAACAAGGCTTGA	1441						
Qy	1441	AGCTTAA	GGCTGTGAGACGCTTCAAGTCTGGGCCCATCTATCAGAGGCGAGGCTTAC	1501	1441	AGCTTAA	GGCTGTGAGACGCTTCAAGTCTGGGCCCATCTATCAGAGGCGAGGCTTAC	1501								
Db	1441	AGCTTAA	GGCTGTGAGACGCTTCAAGTCTGGGCCCATCTATCAGAGGCGAGGCTTAC	1501	1441	AGCTTAA	GGCTGTGAGACGCTTCAAGTCTGGGCCCATCTATCAGAGGCGAGGCTTAC	1501								
Qy	1501	GTGGCCCA	CACAGCGCCCTCAGGCCCATCTCCATGTTCTTCCCGCTAAGACCATCAGCGC	1561	1501	GTGGCCCA	CACAGCGCCCTCAGGCCCATCTCCATGTTCTTCCCGCTAAGACCATCAGCGC	1561								
Db	1501	GTGGCCCA	CACAGCGCCCTCAGGCCCATCTCCATGTTCTTCCCGCTAAGACCATCAGCGC	1561	1501	GTGGCCCA	CACAGCGCCCTCAGGCCCATCTCCATGTTCTTCCCGCTAAGACCATCAGCGC	1561								
Qy	1561	CGTCAAA	AGTTTCA	CAGGTGTCACAGGCATNAGACAAAGA	CAAAAGCTTTAAAGCTGTGA	1621	1561	CGTCAAA	AGTTTCA	CAGGTGTCACAGGCATNAGACAAAGA	CAAAAGCTTTAAAGCTGTGA	1621				
Db	1561	CGTCAAA	AGTTTCA	CAGGTGTCACAGGCATNAGACAAAGA	CAAAAGCTTTAAAGCTGTGA	1621	1561	CGTCAAA	AGTTTCA	CAGGTGTCACAGGCATNAGACAAAGA	CAAAAGCTTTAAAGCTGTGA	1621				
Qy	1621	GTTCGGGG	CCAAAGAAAGTTTGAATTTGCTCTCAGAGAGCAGTGAGGCGCTTGTGAGC	1681	1621	GTTCGGGG	CCAAAGAAAGTTTGAATTTGCTCTCAGAGAGCAGTGAGGCGCTTGTGAGC	1681								
Db	1621	GTTCGGGG	CCAAAGAAAGTTTGAATTTGCTCTCAGAGAGCAGTGAGGCGCTTGTGAGC	1681	1621	GTTCGGGG	CCAAAGAAAGTTTGAATTTGCTCTCAGAGAGCAGTGAGGCGCTTGTGAGC	1681								
Qy	1681	ACCCAGA	AGTATCTCAAGTGAAGAGAAAC	CTGTGA	GTTTAACTTACCTGACGATATGCGAG	1741	1681	ACCCAGA	AGTATCTCAAGTGAAGAGAAAC	CTGTGA	GTTTAACTTACCTGACGATATGCGAG	1741				
Db	1681	ACCCAGA	AGTATCTCAAGTGAAGAGAAAC	CTGTGA	GTTTAACTTACCTGACGATATGCGAG	1741	1681	ACCCAGA	AGTATCTCAAGTGAAGAGAAAC	CTGTGA	GTTTAACTTACCTGACGATATGCGAG	1741				
Qy	1741	AGATCCC	CGAAATATCA	CCTCAAAAGAACCTTTGGAACAAATACCAACCA	CATACACACTA	1801	1741	AGATCCC	CGAAATATCA	CCTCAAAAGAACCTTTGGAACAAATACCAACCA	CATACACACTA	1801				
Db	1741	AGATCCC	CGAAATATCA	CCTCAAAAGAACCTTTGGAACAAATACCAACCA	CATACACACTA	1801	1741	AGATCCC	CGAAATATCA	CCTCAAAAGAACCTTTGGAACAAATACCAACCA	CATACACACTA	1801				
Qy	1801	CACTCAAA	AGATCA	CATGATTCCTTATTTGGGCTTTGGA	AAAA	CAAGGATGAAGCA	ATTCTCT	1861	1801	CACTCAAA	AGATCA	CATGATTCCTTATTTGGGCTTTGGA	AAAA	CAAGGATGAAGCA	ATTCTCT	1861
Db	1801	CACTCAAA	AGATCA	CATGATTCCTTATTTGGGCTTTGGA	AAAA	CAAGGATGAAGCA	ATTCTCT	1861	1801	CACTCAAA	AGATCA	CATGATTCCTTATTTGGGCTTTGGA	AAAA	CAAGGATGAAGCA	ATTCTCT	1861
Qy	1861	AACTGCT	CTCTAATGGGAGTGGCTTG	CGCAGCA	AGGTCTCACTGTGTATAC	CAACGACA	1921	1861	AACTGCT	CTCTAATGGGAGTGGCTTG	CGCAGCA	AGGTCTCACTGTGTATAC	CAACGACA	1921		
Db	1861	AACTGCT	CTCTAATGGGAGTGGCTTG	CGCAGCA	AG											

Qy	971	-----	970
Db	1021	ATGGCCAGCAGCTGCTTGAGACGAGATGCAGTGTACAGAAAGAAAGTTCACCGGTTAG	1080
Qy	971	-----	970
Db	1081	AAAGCAGCCAGCGGTGGCGCACCTGTATCCCACTACTCGGAGGCTGAGCGAG	1140
Qy	971	-----GGTGGCAGAGCAGCTCATCAACCCCTT	997
Db	1141	GAGATCGCTTGAAACCCGGGAGCGGAGGTTGTGGTGGCAGAGCAGCTCATCAACCCCTT	1200
Qy	998	TGGAGAGATGATGATGATTTTGGAGCACACTGATTTTCGACAGGAATTTGACAGTTC	105
Db	1201	TGGAGAGATGATGATGATTTTGGAGCACACTGATTTTCGACAGGAATTTGACAGTTC	126
Qy	1058	CCGTGTGGCTGTGATGATGATGACCAAGAACCTGCTCGGATGAGCCGAGCATGTACTG	1117
Db	1261	CCGTGTGGCTGTGATGATGATGACCAAGAACCTGCTCGGATGAGCCGAGCATGTACTG	1320
Qy	1118	GATAAGCCGAGCCAGCCAGCCCCCTCAACACCTGTCTTCGCGCCAGTTCCGTCAGCCCTC	1177
Db	1321	GATTAAGCCGAGCCAGCCAGCCCCCTCAACACCTGTCTTCGCGCCAGTTCCGTCAGCCCTC	1380
Qy	1178	CTTTATGGGCTTCACCTTCAACATGACCTGTGAACAAAGAGAGATGATGATTCAGCCCAA	1237
Db	1381	CTTTATGGGCTTCACCTTCAACATGACCTGTGAACAAAGAGAGATGATGATTCAGCCCAA	1440
Qy	1238	TCAGGAGGACGAGGAGGATGCTCAAGCTGGGATTCATTTGGCCGCTTCTTAGGCTTCGACCTC	1297
Db	1441	TCAGGAGGACGAGGAGGATGCTCAAGCTGGGATTCATTTGGCCGCTTCTTAGGCTTCGACCTC	1500
Qy	1298	CCATGATCAACATCTCTCCAGGGCAAACTCAAGACCAAACTACTGTGGCCCAAGAGGGA	1357
Db	1501	CCATGATCAACATCTCTCCAGGGCAAACTCAAGACCAAACTACTGTGGCCCAAGAGGGA	1560
Qy	1358	ATCCCTTCCACGAGGGGCTTGCCCAAAACCAAGGACGACCAAGAAAGCTTAAAGGGG	1417
Db	1561	ATCCCTTCCACGAGGGGCTTGCCCAAAACCAAGGACGACCAAGAAAGCTTAAAGGGG	1620
Qy	1418	CCAGGAGACAAACAAGGCTCTGAAAGCTTAAAGGCTGTGACGCTTCAAGTGTGGCCCACT	1477
Db	1621	CCAGGAGACAAACAAGGCTCTGAAAGCTTAAAGGCTGTGACGCTTCAAGTGTGGCCCACT	1680
Qy	1478	GTAATCAGAGGCGCAGGCTTAATAAGTGCCTCAAGACGCGCTTCAGCCCACTCCCATGTT	1537
Db	1681	GTAATCAGAGGCGCAGGCTTAATAAGTGCCTCAAGACGCGCTTCAGCCCACTCCCATGTT	1740
Qy	1538	CTTCCCCCTTAAGCCATCAGACGCGGTCAAGACTTCAAGTGTCAAGGCAATAGACCA	1597
Db	1741	CTTCCCCCTTAAGCCATCAGACGCGGTCAAGACTTCAAGTGTCAAGGCAATAGACCA	1800
Qy	1598	AGACAAAGCTTAAAGACTGTAGTTCTGGGCGCAAGAAAGTTTGAATTTGCTCTCAGA	1657
Db	1801	AGACAAAGCTTAAAGACTGTAGTTCTGGGCGCAAGAAAGTTTGAATTTGCTCTCAGA	1860
Qy	1658	GAGGATGGGGCCCTTGATGAGGACCCCAAGAAAGTATCTCAAGTGAAGGAAACCTGTGGA	1717
Db	1861	GAGGATGGGGCCCTTGATGAGGACCCCAAGAAAGTATCTCAAGTGAAGGAAACCTGTGGA	1920
Qy	1718	GTTTAACTTGAACGGATATGCGACAGATCCCGGAAATCACTCAAGAAAGCTTTGGAAAC	1777
Db	1921	GTTTAACTTGAACGGATATGCGACAGATCCCGGAAATCACTCAAGAAAGCTTTGGAAAC	1980
Qy	1778	ATACCAACCAACATATACACTACACTCAAAAGATCACATGATCTTTATGGGCTTTGGA	1837
Db	1981	ATACCAACCAACATATACACTACACTCAAAAGATCACATGATCTTTATGGGCTTTGGA	2040
Qy	1838	AAACAGGATGAAGCACTTCTTAACTGCTTCTTCAATGGGATGCTTGGCCAGCGAGGT	1897
Db	2041	AAACAGGATGAAGCACTTCTTAACTGCTTCTTCAATGGGATGCTTGGCCAGCGAGGT	2100

QY	1898	CCTCAACCTGTGTGAACACACACAGACA	CACTATCTCAAGCAAGCCATACACTGTCCACA	1957
Dp	2101	CCTCACTGTGTGTATACACACAGACA	CACTATCTCAAGTCACAGCCATACACTGTCCACA	2160
QY	1958	CTGAAGAAACCGTGTCTTACACAGCCCTGAATCAAT	TGGTTAGCTTAATAGATAAAAATCCC	2017
Dp	2161	CTGAAGAAACCGTGTCTTACACAGCCCTGAATCAAT	TGGTTAGCTTAATAGATAAAAATCCC	2220
QY	2018	AGACTACTTCAGCCCTTAAATGCTTTATTCAT	TAAATACGTGGAACCTAGACTGAACCA	2077
Dp	2221	AGACTACTTCAGCCCTTAAATGCTTTATTCAT	TAAATACGTGGAACCTAGACTGAACCA	2280
QY	2078	TTGGAACAATTTA	CTCAGACTCTCGAATTCAGAGTCGGAAACCTTAGTTCTATCTGAAT	2137
Dp	2281	TTGGAACAATTTA	CTCAGACTCTCGAATTCAGAGTCGGAAACCTTAGTTCTATCTGAAT	2340
QY	2138	CCAAGACAGCCACA	CCTTAGATACTGCCCAAACTAATGAGTTAATTAATACAAATACT	2197
Dp	2341	CCAAGACAGCCACA	CCTTAGATACTGCCCAAACTAATGAGTTAATTAATACAAATACT	2400
QY	2198	CGTTAAAAA	AAAAAAAAAAAAAAAAAAAAA	2226
Dp	2401	CGTTAAAAA	AAAAAAAAAAAAAAAAAAAAA	2429

RESULT 3
 AD084435
 AD084435 standard; cDNA; 2404 BP.
 ID
 AC AD084435;
 XX
 DT 07-OCT-2004 (first entry)
 XX
 DE Human tumour-associated antigenic target (TAT) cDNA sequence #1249.
 XX
 KM human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;
 XX cancer; cell proliferative disorder; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO2004060270-A2.
 XX
 PD 22-JUL-2004.
 XX
 PF 15-OCT-2003; 2003WO-US029126.
 XX
 PR 18-OCT-2002; 2002US-0418988P.
 XX
 PA (GETH) GENENTECH INC.
 XX (WUTD/) WU T D.
 XX (ZHOU/) ZHOU Y.
 PI
 XX Wu TD, Zhou Y;
 XX
 DR WPI; 2004-534300/51.
 XX
 PT New nucleic acid molecule and encoded polypeptide, for diagnosing,
 PT preventing or treating cell proliferative disorders such as cancer.
 XX
 PS Claim 1; SEQ ID NO 1249; 5504bp; English.
 CC
 CC The present invention describes an isolated tumour-associated antigenic
 CC target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide
 CC sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of
 CC (a); (c) the complement of (a) or (b); (d) a sequence that has 80%
 CC sequence identity to (a) -(c); or (e) a sequence that hybridises to (a) -
 CC (c). Also described: (1) an expression vector comprising the above
 CC nucleic acid; (2) a host cell comprising the above expression vector; (3)
 CC a process for producing a polypeptide; (4) an isolated polypeptide
 CC comprising: (a) an amino acid sequence encoded by any of the above
 CC nucleotide sequences; (b) an amino acid sequence encoded by the full-
 CC length coding region of the above nucleotide sequences; or (c) a sequence
 CC having at least 80% identical to (a) or (b); (5) a chimeric polypeptide

CC comprising the above polypeptide fused to a heterologous polypeptide; (6)
CC an isolated antibody that binds to the above polypeptide; (7) a process
CC for producing the antibody; (8) an isolated oligopeptide that binds to
CC the above polypeptide; (9) a tumour-associated antigenic target (TAT)
CC binding organic molecule that binds to the above polypeptide; (10) a
CC composition of matter comprising the above (chimeric) polypeptide,
CC antibody, oligopeptide or TAT binding organic molecule, in combination
CC with a carrier; (11) an article of manufacture comprising a container and
CC the composition of matter contained within the container; (12) methods of
CC inhibiting the growth of a cell that expresses the above protein, where
CC the growth of the cell is at least in part dependent upon a growth
CC potentializing effect of the above protein; (13) a method of
CC therapeutically treating a mammal having a cancerous tumour comprising
CC cells that express the above protein; (14) a method of determining the
CC presence of a protein in a sample suspected of containing the protein
CC described above; (15) methods of diagnosing the presence of a tumour in a
CC mammal; (16) a method for treating or preventing a cell proliferative
CC disorder associated with increased expression or activity of the above
CC protein; and (17) a method of binding an antibody, oligopeptide or
CC organic molecule to a cell that expresses the protein described above.
CC The TAT sequences have cytostatic activities, and can be used in gene
CC therapy. The composition and methods are useful for diagnosing,
CC preventing or treating cancer. The composition is also used for preparing
CC a medicament for the therapeutic treatment or diagnostic detection of a
CC cell proliferative disorder or cancer. The present sequence represents a
CC human TAT cDNA sequence from the present invention.

XX Sequence 2404 BP; 590 A; 696 C; 611 G; 507 T; 0 U; 0 Other;

Query Match 89.0%; Score 1984.8; DB 12; Length 2404;
Best Local Similarity 91.5%; Pred. No. 0;
Matches 2199; Conservative 0; Mismatches 2; Indels 203; Gaps 1;

QY 1 CAGGAGATCCCAACAGCTTAATGCGCAGACCTTCTGTGGGATCATCGGACCCACTTGGAA 60
DB 1 CAGGAGATCCCAACAGCTTAATGCGCAGACCTTCTGTGGGATCATCGGACCCACTTGGAA 60
QY 61 CCCACCTGACCCCAAGCCCACTCTGTGAGCCCACTGGCTGGCATGACCATCACTTACA 120
DB 61 CCCACCTGACCCCAAGCCCACTCTGTGAGCCCACTGGCTGGCATGACCATCACTTACA 120
QY 121 CAAGCCAAAGTGGCTAATGCCCGCTTAAGGCTCTTCTCCCGCTGTGCTGTGCTGGCGGG 180
DB 121 CAAGCCAAAGTGGCTAATGCCCGCTTAAGGCTCTTCTCCCGCTGTGCTGTGCTGGCGGG 180
QY 121 CAAGCCAAAGTGGCTAATGCCCGCTTAAGGCTCTTCTCCCGCTGTGCTGTGCTGGCGGG 180
DB 121 CAAGCCAAAGTGGCTAATGCCCGCTTAAGGCTCTTCTCCCGCTGTGCTGTGCTGGCGGG 180
QY 181 GCAGCATCTACAGCTGTATATGCGAGTCTTAATCTTCTGCTCTGACTACTACATCA 240
DB 181 GCAGCATCTACAGCTGTATATGCGAGTCTTAATCTTCTGCTCTGACTACTACATCA 240
QY 241 TCCGCTTATTTATTAAGGCTGGCCCTCA CGGAAGAACAAAGCTGATGTTGAGAAACGA 300
DB 241 TCCGCTTATTTATTAAGGCTGGCCCTCA CGGAAGAACAAAGCTGATGTTGAGAAACGA 300
QY 301 CTCTGTATTGAGACAGCTACATCCAGCTCATCCCATTTCTCTTGTGTGCTGCTTCAAG 360
DB 301 CTCTGTATTGAGACAGCTACATCCAGCTCATCCCATTTCTCTTGTGTGCTGCTTCAAG 360
QY 361 TGACGCTGTGTGTACCGCGCTGTGTGAACAGTACGAGAACTTGCCTGTCGACCGCC 420
DB 361 TGACGCTGTGTGTACCGCGCTGTGTGAACAGTACGAGAACTTGCCTGTCGACCGCC 420
QY 421 TCATGAGCTGTGTGTGCGGCGCTTGTGCGAAGGCAAGAGCGAGGCGGCTGTGGGGC 480
DB 421 TCATGAGCTGTGTGTGCGGCGCTTGTGCGAAGGCAAGAGCGAGGCGGCTGTGGGGC 480
QY 481 GCAAGCTCATTCGCTAGCGCAACCTGGGCAACGTGTCTATCTCTGCGAGCTGACGACCG 540
DB 481 GCAAGCTCATTCGCTAGCGCAACCTGGGCAACGTGTCTATCTCTGCGAGCTGACGACCG 540
QY 541 CAGGCTACAGGCGCTTCCCGACGCGCCAGCACTGTGTGCAAGGCTTTATGACTCGGG 600
DB 541 CAGGCTACAGGCGCTTCCCGACGCGCCAGCACTGTGTGCAAGGCTTTATGACTCGGG 600

QY 601 CAGAACCAAGAGCTTGAAGAACTGAGCCCTACCAACAACATGTTGTGGTGCCTGGG 660
DB 601 CAGAACCAAGAGCTTGAAGAACTGAGCCCTACCAACAACATGTTGTGGTGCCTGGG 660
QY 661 TGTGTTTGGCAACCTTCAATGAAGGCGTGTGTGAAGGTGCAATCCGGGACCTTATCC 720
DB 661 TGTGTTTGGCAACCTTCAATGAAGGCGTGTGTGAAGGTGCAATCCGGGACCTTATCC 720
QY 721 TGTCCAGAGCTGTGCTGAAGAGATGAACCTTGGCTACTGATGTGACACCTGTATG 780
DB 721 TGTCCAGAGCTGTGCTGAAGAGATGAACCTTGGCTACTGATGTGACACCTGTATG 780
QY 781 CTTACGACTGATATGATATCCACTGTGTATACAGAGTGTGACTGTGGCGGTGACA 840
DB 781 CTTACGACTGATATGATATCCACTGTGTATACAGAGTGTGACTGTGGCGGTGACA 840
QY 841 GCTTCTTCTGACTTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900
DB 841 GCTTCTTCTGACTTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900
QY 901 GCCATGAGCTGAGCTGT 960
DB 901 GCCATGAGCTGAGCTGT 960
QY 961 GCTGGCTGAA----- 970
DB 961 GCTGGCTGAAAGTGTGGGCTCTCCAGGGCCGTGCTGTGAGGATGAGGAGGAGG 1020
QY 971 ----- 970
DB 1021 ATGCGCAGAGCTGCTGAGACGAGATGAGATGAGTGTGAGAAAGAGTCTCACGGTATG 1080
QY 971 ----- 970
DB 1081 AAGACAGCCAGGCGT 1140
QY 971 -----GTTGCAAGAGCTGATCAACCCCTT 997
DB 1141 GAGAAATGCTTGAACCGGGAGGCGAGGATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1200
QY 998 TGGAGAGATGATGATATTTTGAACCACTGTGATGTGTGACAGAAATTTGCAAGGTTC 1057
DB 1201 TGGAGAGATGATGATATTTTGAACCACTGTGATGTGTGACAGAAATTTGCAAGGTTC 1260
QY 1058 CCGT 1117
DB 1261 CCGT 1320
QY 1118 GAATTAAGCCGAGGACAGGCCCCCTTACAGAGCTGTTCGCGCAGTTCCGTGAGGCTC 1177
DB 1321 GAATTAAGCCGAGGACAGGCCCCCTTACAGAGCTGTTCGCGCAGTTCCGTGAGGCTC 1380
QY 1178 CTTTATGGGCTCACTTCAACATGAGCTGTGAACAAAGAGAGATGAGTTCAAGCCAA 1237
DB 1381 CTTTATGGGCTCACTTCAACATGAGCTGTGAACAAAGAGAGATGAGTTCAAGCCAA 1440
QY 1238 TCAGAGAGAGAGAGAGATGCTACAGCTGTGCAATATTGGCGCGCTTCTTAAGGCTGCAATC 1297
DB 1441 TCAGAGAGAGAGAGAGATGCTACAGCTGTGCAATATTGGCGCGCTTCTTAAGGCTGCAATC 1500
QY 1298 CCAATGATCAACATCTCTCCAGAGGCAAACTCAAGGACCAAACTACTGTGCGCAAGAGGA 1357
DB 1501 CCAATGATCAACATCTCTCCAGAGGCAAACTCAAGGACCAAACTACTGTGCGCAAGAGGA 1560
QY 1358 ATCCCTTCTTCAAGAGGCTGTGCCAAAAAACAAGAGGACCAAAAGAAAGTTAGGGG 1417
DB 1561 ATCCCTTCTTCAAGAGGCTGTGCCAAAAAACAAGAGGACCAAAAGAAAGTTAGGGG 1620
QY 1418 CCAAG 1477
DB 1621 CCAAG 1680
QY 1478 GTATCAAGAGGCGAGGCTACTACAGTGGCCCAAGACCGCCCTCAGCCCACTCCCATGTT 1537

```

Db      1681  GTATCAGAGGCGAGGCTACTACAGTGCCCAAGAGGCCCTCAGGCCCACTCCATGTT 1740
Qy      1538  CTTCCCGCCAGAACCATCAGCCCGCTCAAAAGCTTCACTGTCACAGGATAGACCA 1597
Db      1741  CTTCCCGCCAGAACCATCAGCCCGCTCAAAAGCTTCACTGTCACAGGATAGACCA 1800
Qy      1538  AGACAAAAGCTTAAAGACTGAGTTCTGGGSCCAAGAAAAGTTTGAATTGCTCTCA 1657
Db      1801  AGACAAAAGCTTAAAGACTGAGTTCTGGGSCCAAGAAAAGTTTGAATTGCTCTCA 1860
Qy      1658  GAGCGATGGGCGCTTGAATGAGCACCAGAAATATCTCAAGTGAAGAGAAAAGTGTGA 1717
Db      1861  GAGCGATGGGCGCTTGAATGAGCACCAGAAATATCTCAAGTGAAGAGAAAAGTGTGA 1920
Qy      1718  GTTAAACCTGACGGATATGCCAGAGATCCCGAAAATACCTCAAGAAAGCTTTGAA 1777
Db      1921  GTTAAACCTGACGGATATGCCAGAGATCCCGAAAATACCTCAAGAAAGCTTTGAA 1980
Qy      1778  ATCAACCAACCAATACACACTCACTCAAAAGATCAATGATCTTATTTGGGCTTGA 1837
Db      1981  ATCAACCAACCAATACACACTCACTCAAAAGATCAATGATCTTATTTGGGCTTGA 2040
Qy      1838  AAACAGGATGAGACACTTCTTAACTGCTTCTTAATGGGAGTCTTCCGACGAGGT 1897
Db      2041  AAACAGGATGAGACACTTCTTAACTGCTTCTTAATGGGAGTCTTCCGACGAGGT 2100
Qy      1898  CCTCACTGTGTGTACACAGAGAGACACTGATCCAGTCAAGCAGCATACAGCTGTCA 1957
Db      2101  CCTCACTGTGTGTACACAGAGAGACACTGATCCAGTCAAGCAGCATACAGCTGTCA 2160
Qy      1958  CTGAAGAACTGTCTCTTACACAGCCTGATCAAAATGTTAGCTTAATGATAAAAATCCC 2017
Db      2161  CTGAAGAACTGTCTCTTACACAGCCTGATCAAAATGTTAGCTTAATGATAAAAATCCC 2220
Qy      2018  AGACTACTTCAAGCCTTTAATGCTTTTATATATAAACTGTGAAAGCTAGACTGAACA 2077
Db      2221  AGACTACTTCAAGCCTTTAATGCTTTTATATATAAACTGTGAAAGCTAGACTGAACA 2280
Qy      2078  TTGAAACATTTAACTCAGACTGATCTGATTCAGAGTCGGGAACCTTATGTTCTATCTG 2137
Db      2281  TTGAAACATTTAACTCAGACTGATTCAGAGTCGGGAACCTTATGTTCTATCTGAT 2340
Qy      2138  CCAAGACAGCCACACTTAGTATCTGCCCAACTAAATGATTTAATAAATACAAATACT 2197
Db      2341  CCAAGACAGCCACACTTAGTATCTGCCCAACTAAATGATTTAATAAATACAAATACT 2400
Qy      2198  CGTT 2201
Db      2401  CGTT 2404

```

RESULT 4
AD083262
ID AD083262 standard; cDNA; 2404 BP.
XX

AD083262;
07-OCT-2004 (first entry)
Human tumour-associated antigenic target (TAT) cDNA sequence #76.
human, tumour-associated antigenic target; TAT; cytostatic; gene therapy;
cancer; cell proliferative disorder; gene; ss.
Homo sapiens.
OS
PN W02004060270-A2.
XX
PD 22-JUL-2004.
XX
PF 15-OCT-2003; 2003MO-US029126.
XX

```

PR      18-OCT-2002; 2002US-0418988P.
XX
XX (GETH ) GENENTECH INC.
PA (W02D ) WU T D.
PA (ZHOU ) ZHOU Y.
XX
PI Wu TD, Zhou Y;
XX
XX WPI; 2004-534300/51.
XX
XX New nucleic acid molecule and encoded polypeptide, for diagnosing,
XX preventing or treating cell proliferative disorders such as cancer.
XX
XX Claim 1; SEQ ID NO 76; 5504pp; English.
XX
XX The present invention describes an isolated tumour-associated antigenic
XX target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide
XX sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of
XX (a); (c) the complement of (a) or (b); (d) a sequence that has 80%
XX sequence identity to (a) -(c); or (e) a sequence that hybridises to (a) -
XX (c). Also described: (1) an expression vector comprising the above
XX nucleic acid; (2) a host cell comprising the above expression vector; (3)
XX a process for producing a polypeptide; (4) an isolated polypeptide
XX comprising: (a) an amino acid sequence encoded by any of the above
XX nucleotide sequences; (b) an amino acid sequence encoded by the full-
XX length coding region of the above nucleotide sequences; or (c) a sequence
XX having at least 80% identical to (a) or (b); (5) a chimeric polypeptide
XX comprising the above polypeptide fused to a heterologous polypeptide; (6)
XX an isolated antibody that binds to the above polypeptide; (7) a process
XX for producing the antibody; (8) an isolated oligopeptide that binds to
XX the above polypeptide; (9) a tumour-associated antigenic target (TAT)
XX binding organic molecule that binds to the above polypeptide; (10) a
XX composition of matter comprising the above (chimeric) polypeptide,
XX antibody, oligopeptide or TAT binding organic molecule, in combination
XX with a carrier; (11) an article of manufacture comprising a container and
XX the composition of matter contained within the container; (12) methods of
XX inhibiting the growth of a cell that expresses the above protein, where
XX the growth of the cell is at least in part dependent upon a growth
XX potentiating effect of the above protein; (13) a method of
XX therapeutically treating a mammal having a cancerous tumour comprising
XX cells that express the above protein; (14) a method of determining the
XX presence of a protein in a sample suspected of containing the protein
XX described above; (15) methods of diagnosing the presence of a tumour in a
XX mammal; (16) a method for treating or preventing a cell proliferative
XX disorder associated with increased expression or activity of the above
XX protein; and (17) a method of binding an antibody, oligopeptide or
XX organic molecule to a cell that expresses the protein described above.
XX The TAT sequences have cytostatic activities, and can be used in gene
XX therapy. The composition and methods are useful for diagnosing,
XX preventing or treating cancer. The composition is also used for preparing
XX a medicament for the therapeutic treatment or diagnostic detection of a
XX cell proliferative disorder or cancer. The present sequence represents a
XX human TAT cDNA sequence from the present invention.
XX
XX Sequence 2404 BP; 590 A; 696 C; 611 G; 507 T; 0 U; 0 Other;
XX
XX Query Match      89.0%; Score 1984.8; DB 13; Length 2404;
XX Best Local Similarity 91.5%; Pred. No. 0;
XX Matches 2199; Conservative 0; Mismatches 2; Indels 203; Gaps 1;

```



```

Db 181 |GCACATCTACAACTGCTATATGCGAGTCTTAATCTTCTGCTGCTACTACATCA 240
Qy 241 |TCCGCTTTATTTAATAGCTGGCCCTCAGCGAAGAAACAACAGCTAATGTTGAAACATGA 300
Db 241 |TCCGCTTTATTTAATAGCTGGCCCTCAGCGAAGAAACAACAGCTAATGTTGAAACATGA 300
Qy 301 |CTCTGTATTTGGCAGAGCTACATCCAGCTCATCCCATTTCTTGGTGGGCTTCTACG 360
Db 301 |CTCTGTATTTGGCAGAGCTACATCCAGCTCATCCCATTTCTTGGTGGGCTTCTACG 360
Qy 361 |TGACGCTGTGTGTGACCCGCTGTGTGAAACAAGTACAGAACCTGCGGTGACCGCC 420
Db 361 |TGACGCTGTGTGTGACCCGCTGTGTGAAACAAGTACAGAACCTGCGGTGACCGCC 420
Qy 421 |TCATGACCTGTGTGTGCGGCTTTCGTGAAAGCAGAGACAGAACCGCGCTGTGTGCGG 480
Db 421 |TCATGACCTGTGTGTGCGGCTTTCGTGAAAGCAGAGACAGAACCGCGCTGTGTGCGG 480
Qy 481 |GCACGCTCATCGGCTACGCAACCTGGGCAACGTGCTCATCTGGGCGAGCGTCAAGCACG 540
Db 481 |GCACGCTCATCGGCTACGCAACCTGGGCAACGTGCTCATCTGGGCGAGCGTCAAGCACG 540
Qy 541 |CAGTCTACAAAGCGCTTCCCAAGCGCCAGCACTGTGTGCAAGCAGGCTTTATGACTCCG 600
Db 541 |CAGTCTACAAAGCGCTTCCCAAGCGCCAGCACTGTGTGCAAGCAGGCTTATGACTCCG 600
Qy 601 |CAGAAACACAAAGAGTTGGAGAAATGAGCCTTACCAACATGTTCTGAGGTGCGCTGGG 660
Db 601 |CAGAAACACAAAGAGTTGGAGAAATGAGCCTTACCAACATGTTCTGAGGTGCGCTGGG 660
Qy 661 |TGTGTTTGGCCAACTGTCAATGAAAGGCGTGGCTTGAAGTCSAATCCGGGACCTTATCC 720
Db 661 |TGTGTTTGGCCAACTGTCAATGAAAGGCGTGGCTTGAAGTCSAATCCGGGACCTTATCC 720
Qy 721 |TGCTCCAGAGGCTCTGTAACAGATGAAACAACCTTGCCTACTCAGTGTGACACCTGTATG 780
Db 721 |TGCTCCAGAGGCTCTGTAACAGATGAAACAACCTTGCCTACTCAGTGTGACACCTGTATG 780
Qy 781 |CCTACGACTGGATTAATGTTCCCACTGTGTATACACAGGTGTGATCTGTGGCGGTGTA 840
Db 781 |CCTACGACTGGATTAATGTTCCCACTGTGTATACACAGGTGTGATCTGTGGCGGTGTA 840
Qy 841 |GCTTCTTCTGCTGCTGTCTAGTTGGGCGGAGTTTCTGAAACCGACGCAAGGCTTACCTG 900
Db 841 |GCTTCTTCTGCTGCTGTCTAGTTGGGCGGAGTTTCTGAAACCGACGCAAGGCTTACCTG 900
Qy 901 |GCCATGAGCTGGACCTGCTGTGTGCGGCTTCAAGTTCTGTGCAAGTCTTCTTATGTTG 960
Db 901 |GCCATGAGCTGGACCTGCTGTGTGCGGCTTCAAGTTCTGTGCAAGTCTTCTTATGTTG 960
Qy 961 |GCTGGCTGGA----- 970
Db 961 |GCTGGCTGGAAGGTGGGCTCTCCAGGCGCTGTGGGCTGGAGGCAATGSCCAGAGGGGTC 1020
Qy 971 |----- 970
Db 971 |----- 970
Qy 1021 |ATGGCCACAGAGCTGCTGAGACAGAGATGACGTCAAGAAAGAAAGTCTTCAAGGGTAG 1080
Db 1021 |ATGGCCACAGAGCTGCTGAGACAGAGATGACGTCAAGAAAGAAAGTCTTCAAGGGTAG 1080
Qy 971 |----- 970
Db 971 |----- 970
Qy 971 |-----GTTGGCAGAGACGCTCATCAACCCCTT 997
Db 1141 |GAGAAATGCTTTGAAACCCGGAAGCGAGGTGTGTGGTGGCAGAGCCTCATCAACCCCTT 1200
Qy 998 |TGGAGAGATGATGATTTTGAAGCACTGGATTGTGACAGGAATTTGGAGGTGTC 1057
Db 1201 |TGGAGAGATGATGATTTTGAAGCACTGGATTGTGACAGGAATTTGGAGGTGTC 1260
Qy 1058 |CTGTGTGCTGTGATGATGACACAGACCTGCTCGATGAGCGGACATGTACTG 1117
Db 1261 |CTGTGTGCTGTGATGATGACACAGACCTGCTCGATGAGCGGACATGTACTG 1130
Qy 1118 |GAATAGCCGAGCAGACAGCCCCCTTACAGAGTGTCTTCCGCGCAGTTCCGTGAGGCTC 1177
Db 1321 |GAATAGCCGAGCAGACAGCCCCCTTACAGAGTGTCTTCCGCGCAGTTCCGTGAGGCTC 1380
Qy 1178 |CTTTATGGGCTTCAACATCAAGCTTGAACAAAGAGAGATGAGATTCCAGCCCA 1237
Db 1381 |CTTTATGGGCTTCAACATCAAGCTTGAACAAAGAGAGATGAGATTCCAGCCCA 1440
Qy 1238 |TCAGAGAGACAGAGAGATGCTCAAGCTGTGCATATTGGCCGCTTCTTAGGCTGCAATC 1297
Db 1441 |TCAGAGAGACAGAGAGATGCTCAAGCTGTGCATATTGGCCGCTTCTTAGGCTGCAATC 1500
Qy 1298 |CCATGATCAACCATCTCCAGGGGCAACTCAAGAACCAATCTACTGTGGCCCAAGAGGA 1357
Db 1501 |CCATGATCAACCATCTCCAGGGGCAACTCAAGAACCAATCTACTGTGGCCCAAGAGGA 1560
Qy 1358 |ATCCCTTCTCCAGAGGCGCTGCCCAAAAACACAGAGCAGCCAAAGAAAGTTAGGGG 1417
Db 1561 |ATCCCTTCTCCAGAGGCGCTGCCCAAAAACACAGAGCAGCCAAAGAAAGTTAGGGG 1620
Qy 1418 |CCAGAAAGACAAAGGCTGTGAAGCTTAAAGCTGTGTGAGCGCTTCAAATCTGGCCCACT 1477
Db 1621 |CCAGAAAGACAAAGGCTGTGAAGCTTAAAGCTGTGTGAGCGCTTCAAATCTGGCCCACT 1680
Qy 1478 |GTATCAGAGGCAAGCTAAGTGTGCCCCAACAAGCCCCCTGAGCCCCACTCCCAATGTT 1537
Db 1681 |GTATCAGAGGCAAGCTAAGTGTGCCCCAACAAGCCCCCTGAGCCCCACTCCCAATGTT 1740
Qy 1538 |CTTCCCTTGAAGACATCAAGGCGGTCAAAAGCTTCAAGTGTCAAGGATAGACACCA 1597
Db 1741 |CTTCCCTTGAAGACATCAAGGCGGTCAAAAGCTTCAAGTGTCAAGGATAGACACCA 1800
Qy 1598 |AGACAAAGCTTAAAGACTGTGAGTGTGTGGGCGCAAGAAAGTTTGAATGCTCTCAGA 1657
Db 1801 |AGACAAAGCTTAAAGACTGTGAGTGTGTGGGCGCAAGAAAGTTTGAATGCTCTCAGA 1860
Qy 1658 |GAGGATGGGCGCTTGAATGAGACCCAGAAAGTATCTCAAGTGAAGAGAAACGTGGA 1717
Db 1861 |GAGGATGGGCGCTTGAATGAGACCCAGAAAGTATCTCAAGTGAAGAGAAACGTGGA 1920
Qy 1718 |GTTTAACTGACGATATGACAGAGATCCCGAAATACACTCAAGAACCTTTGAAACA 1777
Db 1921 |GTTTAACTGACGATATGACAGAGATCCCGAAATACACTCAAGAACCTTTGAAACA 1980
Qy 1778 |ATCACCAACCAACATACACACTCAAGATCAAGATCAAGATCTTATTTGGGCTTGGGA 1837
Db 1981 |ATCACCAACCAACATACACACTCAAGATCAAGATCAAGATCTTATTTGGGCTTGGGA 2040
Qy 1838 |AAAAGAGATGAGACATTTCTTAACTGTCTTCTTAAATGGGAGATGCTTGGCAGCCAGT 1897
Db 2041 |AAAAGAGATGAGACATTTCTTAACTGTCTTCTTAAATGGGAGATGCTTGGCAGCCAGT 2100
Qy 1898 |CCTCACTGTGTGTAACACAGAGAGACATGATCAAGTCAAGCAGCATACAGCTGTCCACA 1957
Db 2101 |CCTCACTGTGTGTAACACAGAGAGACATGATCAAGTCAAGCAGCATACAGCTGTCCACA 2160
Qy 1958 |CTGAAGAACGTGTCTTAAACAGCCTGAATCAAAATGTTAGCTTAAATGAATAAAAATCC 2017
Db 2161 |CTGAAGAACATGTCTTAAACAGCCTGAATCAAAATGTTAGCTTAAATGAATAAAAATCC 2220
Qy 2018 |AGACTACTTCAAGCTTTAAATGCCCTTTAAATCAATAAAACTGTGAAGCTTAAGTGAACA 2077
Db 2221 |AGACTACTTCAAGCTTTAAATGCCCTTTAAATCAATAAAACTGTGAAGCTTAAGTGAACA 2280
Qy 2078 |TTGGAACATTTAACTCAGACTCTGATTCAGAGTCCGGAACCTTATGTTCTATCTGAT 2137
Db 2281 |TTGGAACATTTAACTCAGACTCTGATTCAGAGTCCGGAACCTTATGTTCTATCTGAT 2340
Qy 2138 |CAAAGACGCAACCTTAAATGCTGCCCAAACTAAATGAATTTAAATGAATCAAAATGCT 2197
Db 2341 |CAAAGACGCAACCTTAAATGCTGCCCAAACTAAATGAATTTAAATGAATCAAAATGCT 2400
```



```
QY 1245 GACGAGAGATGCTCAAGCTGCGATCATTTGGCCGCTTCTAGGCTTCAGTCCCATGAT 1304
DB 1141 GACGAGAGATGCTCAAGCTGCGATCATTTGGCCGCTTCTAGGCTTCAGTCCCATGAT 1200
QY 1305 CACCATCTCTCCAGGGGCAAACTCAAGAGCAAACTACTGTGTGGCCCAAGAGGAATCCCTT 1364
DB 1201 CACCATCTCTCCAGGGGCAAACTCAAGAGCAAACTACTGTGTGGCCCAAGAGGAATCCCTT 1260
QY 1365 CTCGACGAGGGCTGGCCCAAAACCAAGAGGAGCAAAAGCTTAAAGGGGCAAGAA 1424
DB 1261 CTCGACGAGGGCTGGCCCAAAACCAAGAGGAGCAAAAGCTTAAAGGGGCAAGAA 1320
QY 1425 GACAAACAGGCTGGAAGCTTAAAGCTGTGAGACGCTTCAAGTCTGAGCCCACTGATCAG 1484
DB 1321 GACAAACAGGCTGGAAGCTTAAAGCTGTGAGACGCTTCAAGTCTGAGCCCACTGATCAG 1380
QY 1485 AGGCGAAGCTACTAAGTGTGCCCCCAAGAGGCGCCCTCAGGCCCACTCCAGTCTTCCCC 1544
DB 1381 AGGCGAAGCTACTAAGTGTGCCCCCAAGAGGCGCCCTCAGGCCCACTCCAGTCTTCCCC 1440
QY 1545 CTAGAACCATCAGGCGCGTCAAAAGCTTCAAGTGTGAGAGGAGTATGACACCAAGACAAA 1604
DB 1441 CTAGAACCATCAGGCGCGTCAAAAGCTTCAAGTGTGAGAGGAGTATGACACCAAGACAAA 1500
QY 1605 AGCTTAAAGACTGTGAGTCTGGGGCCCAAGAAAGTTTGAATTGCTCTCAGAGAGCAT 1664
DB 1501 AGCTTAAAGACTGTGAGTCTGGGGCCCAAGAAAGTTTGAATTGCTCTCAGAGAGCAT 1560
QY 1665 GGGGCTCTTGTATGAGCAAGCAAGATATCTCAAGTGAAGAGAGAAAGTGTGAGTTTAA 1724
DB 1561 GGGGCTCTTGTATGAGCAAGCAAGATATCTCAAGTGAAGAGAGAAAGTGTGAGTTTAA 1620
QY 1725 CTGAGGATATGCAAGAGATCCCGGAAATATCACTCAAGAGAACTTTGGAACATTCACA 1784
DB 1621 CTGAGGATATGCAAGAGATCCCGGAAATATCACTCAAGAGAACTTTGGAACATTCACA 1680
QY 1785 ACCAATATACACTACACTCAAGATATCATGATCTTATTTGGGCTTTGGAAGACAG 1844
DB 1681 ACCAATATACACTACACTCAAGATATCATGATCTTATTTGGGCTTTGGAAGACAG 1740
QY 1845 GATGAAGCAATTCCTAA 1862
DB 1741 GATGAAGCAATTCCTAA 1758

RESULT 6
AAV9722
ID AAV9722 standard; cdna; 1263 BP.
XX
AC AAV9722;
XX
DT 26-APR-1999 (first entry)
XX
DE Human adult retina secreted protein bk112_15 cdna.
XX
KW Secreted protein; human; retina; bk112_15; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 100..885
FT tag= a
XX
XX MO985609-A2.
XX
XX 17-DEC-1998.
XX
XX 08-JUN-1998; 98MO-US011822.
XX
XX 11-JUN-1997; 97US-00873218.
XX
XX 05-JUN-1998; 98US-00092722.
XX
```

```
PA (GEMV ) GENETICS INST INC.
XX
XX Jacobs K, McCoy JM, Lavallee ER, Racine LA, Treacy M, Spaulding V;
PI Agostino MJ, Howes SH, Feichel K;
XX
XX WPI; 1999-080899/07.
XX
XX P-PSDB; AAW95345.
XX
XX New polynucleotides encoding secreted human proteins - derived from human
PT foetal brain, adult testes, foetal kidney, adult thyroid or adult retina
PT cdna libraries.
XX
XX Claim 12; Page 71-72; 113pp; English.
XX
XX This is the nucleotide sequence of cDNA clone bk112_15, which includes an
CC open reading frame for a 261-amino acid polypeptide (see AAW95345). The
CC clone was isolated from a human adult retina cDNA library using methods
CC which are selective for cDNAs encoding secreted proteins, or was
CC identified as encoding a secreted or transmembrane protein on the basis
CC of computer analysis of the amino acid sequence of the encoding protein.
CC Database searches indicate some sequence similarity to known sequences.
CC The invention provides cDNA clones (see AAV9721-33) from human adult
CC thyroid, adult retina, adult testes, foetal kidney and foetal brain that
CC encode novel secreted proteins (see AAW95344-53). Each clone is
CC individually available from deposit clone ATCC 98451 (see also AAV9734-
CC 43). The isolated polynucleotides (PNS) and proteins are predicted to
CC have activities which would make them suitable for treating, preventing
CC or ameliorating medical conditions in humans and animals, although no
CC supporting data is given. Suggested activities include nutritional,
CC cytokine, cell proliferation or differentiation, immune stimulating (e.g.
CC as vaccines) or immune suppressing, haematopoietic regulating, tissue
CC growth, activin/inhibin, chemotactic/chemokinetic, haemostatic,
CC thrombolytic, receptor/ligand, antiinflammatory, cadherin/tumour invasion
CC suppressor, and tumour inhibition activities. The PNS are also stated to
CC be useful for gene therapy
XX
XX
XX Sequence 1263 BP; 397 A; 342 C; 272 G; 252 T; 0 U; 0 Other;
XX
XX Query Match 56.0%; Score 1247.2; DB 2; Length 1263;
XX Best Local Similarity 99.8%; Pred. No. 1e-280;
XX Matches 1249; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX
QY 978 GAGCAGCTCATCAACCCCTTTGAGAGATGATGATTTTGGACCACTGATTTGTC 1037
DB 1 GAGCAGCTCATCAACCCCTTTGAGAGATGATGATTTTGGACCACTGATTTGTC 60
QY 1038 GACAGAAATTTGACGATGTCCTGTTGCTGTGTGATGATGACCAAGAGCTGCTCGG 1097
DB 61 GACAGAAATTTGACGATGTCCTGTTGCTGTGTGATGATGACCAAGAGCTGCTCGG 120
QY 1098 ATGAGCCCGACATGTACTGGAATTAAGCCGAGACAGAGCCCTTCAAGAGCTGCTTCC 1157
DB 121 ATGAGCCCGACATGTACTGGAATTAAGCCGAGACAGAGCCCTTCAAGAGCTGCTTCC 180
QY 1158 GCCAGTTCCGTGAGGCTCTTTATGAGCTTCACCTTCAACATCATGAGCAAGAGAG 1217
DB 181 GCCAGTTCCGTGAGGCTCTTTATGAGCTTCACCTTCAACATCATGAGCAAGAGAG 240
QY 1218 GAGATGAGTTTCCAGGCCAATCAGAGAGCAGAGAGAGATGTCACGCTGATATTGGC 1277
DB 241 GAGATGAGTTTCCAGGCCAATCAGAGAGCAGAGAGAGATGTCACGCTGATATTGGC 300
QY 1278 CGCTTCTAAGGCTCGAGTGTCCATGATCAACATCTCCAGGGGAACTCAAGAGCAAA 1337
DB 301 CGCTTCTAAGGCTCGAGTGTCCATGATCAACATCTCCAGGGGAACTCAAGAGCAAA 360
QY 1338 CTACTGTGCCCCAAGAGAGATCCCTTCTCCACAGAGGCTGCCCCAAGAGCAAGAGCA 1397
DB 361 CTACTGTGCCCCAAGAGAGATCCCTTCTCCACAGAGGCTGCCCCAAGAGCAAGAGCA 420
QY 1398 GCCAAACAGAACTTAAAGGCTCAGAGAGCAAGAGAGCAAGAGCTTAAAGGCTGTGAC 1457
DB 421 GCCAAACAGAACTTAAAGGCTCAGAGAGCAAGAGAGCAAGAGCTTAAAGGCTGTGAC 480
```

QY 1458 GCGTTCAGTCTGCGCCCACTGATCAGAGGCCGCTACTAGTGGCCCAAGACGCC 1517
 DB 481 GCGTTCAGTCTGCGCCCACTGATCAGAGGCCGCTACTAGTGGCCCAAGACGCC 540
 QY 1518 CTCAGAGCCCACTCCCATGTTCTTCCCTAGAACCATCAGGCGGTCAAGACTTCAAGT 1577
 DB 541 CTCAGAGCCCACTCCCATGTTCTTCCCTAGAACCATCAGGCGGTCAAGACTTCAAGT 600
 QY 1578 GTCAAGGCAATAGACACCAAGAGCAAAAGCTTAAAGCTGTGAGTCTTGGGGCCAAAGAA 1637
 DB 601 GTCAAGGCAATAGACACCAAGAGCAAAAGCTTAAAGCTGTGAGTCTTGGGGCCAAAGAA 660
 QY 1638 AGTTTGAATTCCTCTCAGAGAGCGATGGGCGCTTGTAGAGCACCAGAAAGTATCTCAA 1697
 DB 661 AGTTTGAATTCCTCTCAGAGAGCGATGGGCGCTTGTAGAGCACCAGAAAGTATCTCAA 720
 QY 1698 GTGAGAGAGAAACCTGTGAGAGTTTAACTGACGATATGCGAAGATCCCGAAATATCAG 1757
 DB 721 GTGAGAGAGAAACCTGTGAGAGTTTAACTGACGATATGCGAAGATCCCGAAATATCAG 780
 QY 1758 CTCAAGAGACCTTTGGAACATCAACCAACATACACACTACCTCAAGATCATG 1817
 DB 781 CTCAAGAGACCTTTGGAACATCAACCAACATACACACTACCTCAAGATCATG 840
 QY 1818 GATCCTTATTTGGGCTTGGAAAAAGGAGTGAAGCAATTCCTTACCTGCTTCTTAATGG 1877
 DB 841 GATCCTTATTTGGGCTTGGAAAAAGGAGTGAAGCAATTCCTTACCTGCTTCTTAATGG 900
 QY 1878 GGATGCTTGGCGAGCGAGGCTCTCACCTGTGTGTACACAGAGAGACCTGATCCAGTCA 1937
 DB 901 GGATGCTTGGCGAGCGAGGCTCTCACCTGTGTGTACACAGAGAGACCTGATCCAGTCA 960
 QY 1938 CAGCCTATCAGGCTGTCCACATGAGAAAGCTGTCTCAACAGCCTGATCAATGATTA 1997
 DB 961 CAGCCTATCAGGCTGTCCACATGAGAAAGCTGTCTCAACAGCCTGATCAATGATTA 1020
 QY 1998 GCTTAATATGATTAATAATCCAGACTACTTACGCTTATGCTTTATGCTTAAACT 2057
 DB 1021 GCTTAATATGATTAATAATCCAGACTACTTACGCTTATGCTTTATGCTTAAACT 1080
 QY 2058 GTGAAGAGCTAGAGCTGACATGATGGAACATTTAACTGAGACTTGCAGTCCGGA 2117
 DB 1081 GTGAAGAGCTAGAGCTGACATGATGGAACATTTAACTGAGACTTGCAGTCCGGA 1140
 QY 2118 ACCCTTATGCTTATCTGATCCAGAGCAGCAGCAGCTTATGCTTATGCTTAACT 2177
 DB 1141 ACCCTTATGCTTATCTGATCCAGAGCAGCAGCAGCTTATGCTTATGCTTAACT 1200
 QY 2178 GTTAAATTAATTCATTAATCTGCTTAAATTAATTAATTAATTAATTAATTAATTA 2229
 DB 1201 GTTAAATTAATTCATTAATCTGCTTAAATTAATTAATTAATTAATTAATTAATTA 1252

RESULT 7
 ADA44960 ID ADA44960 standard; cDNA; 1238 BP.
 AC ADA44960;
 XX 20-NOV-2003 (first entry)
 DE Human polynucleotide #2.
 KW Human, gene; ssr, genetic disorder; genetic fingerprinting;
 KW autoimmune disorder; multiple sclerosis; systemic lupus erythematosus;
 KW insulin dependent diabetes mellitus; graft-versus-host disease; anaemia;
 KW periodontal disease; bone fracture; cartilage damage;
 KW central nervous system disorder; Alzheimer's disease;
 KW Parkinson's disease;
 KW carbohydrate source; cancer; nutrition; carbon source; nitrogen source;
 KW Homo sapiens.
 OS

XX US2003044935-A1.
 PN 06-MAR-2003.
 PD 21-DEC-2000; 2000US-00746783.
 PF 11-JUN-1997; 97US-0086234P.
 PR 12-JUN-1997; 97US-0086234P.
 PR 08-JUL-1997; 97US-0092115P.
 PR 08-SEP-1997; 97US-0093045P.
 PR 02-OCT-1997; 97US-0093045P.
 PR 27-OCT-1997; 97US-00958304.
 PR 07-NOV-1997; 97US-00958304.
 PR 05-JUN-1998; 98US-00092722.
 PR 11-JUN-1998; 98US-00092722.
 PR 17-JUN-1998; 98US-00092722.
 PR 04-AUG-1998; 98US-00130189.
 PR 08-SEP-1998; 98US-00149633.
 PR 01-OCT-1998; 98US-00165960.
 PR 04-NOV-1998; 98US-00185936.
 XX (JACO/) JACOBS K.
 PA (MCCO/) MCCOY J M.
 PA (LVAL/) LA VALLIE E R.
 PA (COLL/) COLLINS-RACIE L A.
 PA (EVAN/) EVANS C.
 PA (MERB/) MERBERG D.
 PA (TREA/) TREACY M.
 PA (SPAU/) SPAULDING V.
 PI Jacobs K, McCoy JM, La Vallie ER, Collins-Racie LA, Evans C;
 PI Merberg D, Treacy M, Spaulding V;
 PI WPI; 2003-521754/49.
 DR P-PSDB; ADA44961.
 XX New polypeptides and polynucleotides having biological activities, useful
 PT as nutritional sources or supplements, or for treating e.g. autoimmune
 PT diseases, cancers, bone fractures or damages, or central nervous system
 PT disorders.
 PS Claim 12; Page 109; 288pp; English.
 XX The invention relates to human polynucleotides and the polypeptides they
 CC encode. The polynucleotides can be used to express recombinant proteins
 CC for analysis, characterisation or therapeutic use, as markers for tissues
 CC in which the corresponding protein is expressed, as molecular weight
 CC markers on Southern gels, as chromosome markers or tags to identify
 CC chromosomes or to map related gene positions, to compare with endogenous
 CC DNA sequences in patients to identify potential genetic disorders, as
 CC probes to hybridise and discover novel related DNA sequences, as
 CC of information to derive PCR primers for genetic fingerprinting, to raise
 CC anti-protein antibodies and in gene therapy. The proteins can be used to
 CC raise antibodies or to elicit another immune response, as reagents in
 CC assays designed to quantitatively determine levels of the protein in
 CC biological fluids, as markers for tissues in which the corresponding
 CC protein is preferentially expressed and to treat autoimmune disorders
 CC (e.g. multiple sclerosis, systemic lupus erythematosus, insulin dependent
 CC diabetes mellitus or graft-versus-host disease), anaemias, periodontal
 CC disorders (e.g. Alzheimer's disease or Parkinson's disease) and cancers.
 CC The proteins and polynucleotides are also useful as nutritional sources
 CC or supplements, e.g. as carbon, nitrogen or carbohydrate sources. This
 CC sequence represents a human polynucleotide of the invention.
 XX Sequence 1238 BP; 387 A; 336 C; 266 G; 249 T; 0 U; 0 Other;
 SQ
 Query Match 45.4%; Score 1011.2; DB 9; Length 1238;
 Best Local Similarity 97.9%; Pred. 1.3e-225;
 Matches 1226; Conservative 0; Mismatches 3; Indels 23; Gaps 19;
 QY 978 GAGCAGCTCATCAACCCCTTTGAGAGAGATGATGATTTTGGAGACCACTGGAATTGTC 1037

```
Db 1 GAGGAGCTCATCAACCCCTTTGGAGAGGTATGATTTTGAACCAACCTGGATTGTC 60
Qy 1038 GACAGGAATTTGGACAGTGTCCCTGTTGGCTGTGATGAGATGACACAGAGCTGCTCGG 1097
Db 61 GACAGGAATTTGGACAGTGTCCCTGTTGGCTGTGATGAGATGACACAGAGCTGCTCG 119
Qy 1098 ATGGAGCGGAGCATGTACTGGAATAGAGCCGAGCCGACAGCCCCCTTACAGCTGCTCC 1157
Db 120 ATGGAGCGGAGCATGTACTGGAATAGAGCCGAGCCGACAGCCCCCTTACAGCTGCTC 178
Qy 1158 GCCGAGTTCGCTCGAGCTCTCTTTATGGGCTCCACCTTCAACATCAGCTGAAACAAGAG 1217
Db 179 GCCGAGTTCGCTCGAGCTCTCTTTATGGGCTCCACCTTCAACATCAGCTGAAACAAGAG 237
Qy 1218 GAGATGAGATTCCAGCCCAATCAGAGAGACGAGAGAGATGCTCAAGCTGAGCTATGGCC 1277
Db 238 GAGATGAGATTCCAGCCCAATCAGAGAGACGAGAGAGATGCTCAAGCTGAGCTATGGG- 296
Qy 1278 CGCTTCTTAGGCTGAGCTCCATGATACCATCTCCCAAGGGCAAACTCAAGAGCAAA 1337
Db 297 CGCTTCTTAGGCTGAGCTCCATGATACCATCTCCCAAGGGCAAACTCAAGAGCA- 355
Qy 1338 CTACTGTGGCCCAAGAGGGAATCCCTTCTCCAAGAGGCTGCCCCAAAACCAAGAGCA 1397
Db 356 CTACTGTGGCCCAAGAGGGAATCCCTTCTCCAAGAGGCTGCCCCAAAACCAAGAGC- 414
Qy 1398 GCCAAACAGAACTTTAGGAGGAGAGAGACAAAGGCTGAGCTTATAGGCTGTGAC 1457
Db 415 GCCAAACAGAACTTTAGGAGGAGAGAGACAAAGGCTGAGCTTATAGGCTGTGAG- 473
Qy 1458 GCCTTCAGTCTGGGCCCTGTATAGAGGCCAGGCTCTACAGTCCCAAGAGCGCC 1517
Db 474 GCCTTCAGTCTGGGCCCTGTATAGAGGCCAGGCTCTACAGTCCCAAGAGCG- 532
Qy 1518 CTCAGCCCCACATCCCATGTTCTTCCCTAGAACCATGAGCGCTCAAGCTTACAGT 1577
Db 533 CTCAGCCCCACATCCCATGTTCTTCCCTAGAACCATGAGCGCTCAAGCTTACAG- 591
Qy 1578 GTCACAGCATAGACACAAAGAACTTAAAGCTGTGAGTTCTTGGGAGCAAGAA 1637
Db 592 GTCACAGCATAGACACAAAGAACTTAAAGCTGTGAGTTCTTGGGAGCAAG- 650
Qy 1638 AGTTTGAATTTGCTCTCAGAGAGCGATGGGGCTTGTATGAGACCCGAAATCTCA 1697
Db 651 AGTTTGAATTTGCTCTCAGAGAGCGATGGGGCTTGTATGAGACCCGAAATCTC- 709
Qy 1698 GTGAGAGAGAAATCTGTGAAGTTTAACTGACGAGATATGCGAGAGATCCCGAAATGAC 1757
Db 710 GTGAGAGAGAAATCTGTGAAGTTTAACTGACGAGATATGCGAGAGATCCCGAAATGAC- 768
Qy 1758 CTCAAAGAACTTTGGAACATCACCAACCAACATACACTCACTCAAAAGATCACATG 1817
Db 769 CTCAAAGAACTTTGGAACATCACCAACCAACATACACTCACTCAAAAGATCACAT- 827
Qy 1818 GATCCTTTTGGGCTTTGGAACAGGAGATGAAAGACATTTCTTAACTGCTTCCATAGG 1877
Db 828 GATCCTTTTGGGCTTTGGAACAGGAGATGAAAGACATTTCTTAACTGCTTCCATAG- 886
Qy 1878 GGATGCTTCGCGAGCAGAGTCTCTCACTGTGTGTACACAGAGAGACATGATCCAGTCA 1937
Db 887 GGATGCTTCGCGAGCAGAGTCTCTCACTGTGTGTGTACACAGAGAGACATGATCCAGTCA- 945
Qy 1938 CAGCATACAGCTGTCCACTGAAGAAAGTGTCTTACCAACAGCTGATCAATGATGTTA 1997
Db 946 CAGCATACAGCTGTCCACTGAAGAAAGTGTCTTACCAACAGCTGATCAATGATG- 1003
Qy 1998 GCTTATATGATTAATATCCCAAGATCTTCACTTCAAGCTTATATGCTTATATTAATACT 2057
Db 1004 GCTTATATGATTAATATCCCAAGATCTTCACTTCAAGCTTATATGCTTATATTAATA- 1061
Qy 2058 GTGAAAGCTAGACTGAACCATTTGAAACATTTAACTCAGACTCTGATTCAGAGTGGGGA 2117
```

```
Db 1062 GTGAAAGCTAGACTGAACCATTTGAAACATTTAACTCAGACTCTGATTCAGAGTGG- 1119
Qy 2118 ACCCTAGTCTATCTGAATCCAGAGACGACACCTTATGATATGCTGCCAAACTATGA 2177
Db 1120 ACCCTAGTCTATCTGAATCCAGAGACGACACCTTATGATATGCTGCCAAACTAT- 1177
Qy 2178 GTTATATATATCAAAATCTGTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAA 2229
Db 1178 GTTATATATATCAAAATCTGTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAA 1229

RESULT 8
AAZ21229
ID AAZ21229 standard; cDNA; 1916 BP.
XX
AC AAZ21229;
XX
DT 22-NOV-1999 (first entry)
XX
DE Mouse CG1CE cDNA sequence.
XX
KW CG1CE; Best's macular dystrophy; mutation; diagnosis; detection; BMD;
KM age-related macular dystrophy; ss.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT CDS 11..1666
FT FT /*tag= a
FT FT /product= "CG1CE protein"
XX
PN WO9943695-A1.
XX
PD 02-SEP-1999.
XX
PF 22-FEB-1999; 99WO-US003790.
XX
PR 25-FEB-1998; 98US-0075941P.
PR 18-DEC-1998; 98US-0112926P.
XX
PA (MERI ) MERCK & CO INC.
PA (UYUP-) UNIV UPPSALA.
XX
PI Petukhin K, Caskey CT, Metzker M, Wadelius C;
XX
DR WPI; 1999-540560/45.
DR P-PSDB; AAY29955.
XX
XX Human and mouse polynucleotides encoding CG1CE polypeptides.
PS Claim 2; Fig 8; 67bp; English.
CC The present sequence represents the mouse CG1CE cDNA sequence, which when
CC mutated is responsible for Best's macular dystrophy (BMD).
CC Polynucleotides encoding CG1CE are useful for diagnosing whether a
CC patient carries a mutation in the CG1CE gene. Normal and mutated CG1CE
CC proteins are useful for identifying activators and/or inhibitors of these
CC proteins, in order to treat BMD. The CG1CE gene offers a simpler and
CC cheaper method of diagnosing BMD without the need for the presence of the
CC patient. The gene may also be useful to discovering the genetic cause of
CC age-related macular dystrophy
XX
SQ Sequence 1916 BP; 514 A; 499 C; 457 G; 446 T; 0 U; 0 Other;

Query Match 39.1%; Score 871.2; DB 2; Length 1916;
Best Local Similarity 77.1%; Pred. No. 7.2e-193;
Matches 116; Conservative 0; Mismatches 293; Indels 39; Gaps 3;

Qy 96 TGCCTGGCAGTACATCACTTACCAAGCAAGTGGCTAATGCCCGTTAGGCTCTTC 155
Db 2 TGCCAGCAGTACATCACTTACCAAGCAAGTGGCTAATGCCCGTTAGGCTCTTC 61
Qy 156 TCCGCGCTGCTGTGCTGGGAGCATTTCAAGCTGTATATGCGAGATTCTTA 215
```

Db 62 TCTCTCCTCTCTCTGTGCTGCGAGGACCACTCTCAAGCTGCTGTATGAGAAATTCCTT 122
 Qy 216 ATCTTCCTGCTCTGTACTTACATCAATCCGCTTTATTTATAGCGTGGCCCTACAGAAAGA 275
 Db 122 GTCTTCATATTCCTCTACTATTCAATCCGTGGACCTCTACAAATATGTTCTTCCAGTGAT 181
 Qy 276 CAACACCTGATGTTTGGAAATACTGACTCTGTATTGCGACAGCTACATCCAGCTATCCC 335
 Db 182 CAGACACTTGTGTTGAGAAAGCTGCTGTGATCTCGACAGACCTACATTCACCTATCCCT 241
 Qy 336 ATTTCTCTTGCTGGCTGGCTTCTACCTGTGACGCTGTGCGAGCCGCTGTGAAACCACTAC 395
 Db 242 AATATCCCTTGCTCTGGATTCTGATTTGATTTGATTTGGTGTGAGCCCTGTGTGAGCACTAC 301
 Qy 396 GAGAACCTGCGGTGGCCCGACCGCTCATGTAGCCGTGTGCGGCTTTCGTGAAAGCAAG 455
 Db 302 GAGAACTTGGCGGTGGCCCGACCGCTCATGTATCCAGGTGTCTAGCTTCTGTGAGGGCAAG 361
 Qy 456 GACGAGCAAGGCGCGCTGCTGCGCGACGCTCATTCGCTAACGCAACCTGGGCAACGTG 515
 Db 362 GATGAGAAAGCCCGTTGGCTGCGCGACGCTCATCCGCTAACGCAATCTGGGCAAGTG 421
 Qy 516 CTCAATCCGCGCACGCTCAGACCGCAGCTTACAGGCGCTTCCCAAGCCCGACGACCTG 575
 Db 422 CTCAATCCGCGCACATCAGACCTTCGCTCTACAGAGGCTTTCCTCCATCTTACACACTG 481
 Qy 576 GTGCAAGCAGGCTTTATGACTCCGCGCGAAGACACAAGCAGTTGGAAATACTGAGCTTACA 635
 Db 482 GTGCTAGCAGGTTTATGACCCCATGGGAAACATTAAGCAGTTGAGAAAGTTGGACCTTACA 541
 Qy 636 CACAACATGTTCTGGGTGCTGCTGGTGTGGTTTGCCAACTGTCAATGAAAGCGTGGCT 695
 Db 542 CACAACATTTCTGGGTGCTGCTGGTGTGGTTTGCCAACTGTCAATGAAAGCGTATCTT 601
 Qy 696 GAAAGTGAATCCGCGGACCTTATCTCGTGTCAAGAGCGTCTGTAAGAGATGAACCTTG 755
 Db 602 GGAAGTGAATCCGCGGACCGCTCTGCTCAAGCGCTGATTAAGAGGTGTGATCTTTG 661
 Qy 756 CGTACTCAGTGTGGAACACTGTATGCTTACGACTGAGATTAGTATCCCATGCTGTATACA 815
 Db 662 CGTACTCAGTGTGGAACGCTGTATGCTTACGACTGAGATTAGTATCCCATGCTGTATACA 721
 Qy 816 CAGTGTGATCTGTGCGGCTGTACAGCTTCTTCGACTGTGTCTGATTTGGCGGCAAGTT 875
 Db 722 CAGTGTGTACAGTGGCAATATACAGCTTTTCTTGACATGCTGTATCGGAGGCAAGTT 781
 Qy 876 CTGAATCCGACCAAGGCTTACCTGTGGCAATGACTGAGCTGAGCTGTTGGTGGCCGTTACG 935
 Db 782 CTGAATCCCAAAACAAGACTTACCAAGGCATGAGATGATCTGTGTGTGCCGTCTTACA 841
 Qy 936 TTCTGCAAGTCTTCTTCTATGTGTGGCGCTGAAAGGTGTGAGAGCAAGCTCATCAACCC 995
 Db 842 ATCTGCAATTTCTTATTTCTACATGCGCTGTGAAAGGTGTGAGAGCAAGCTCATCAACCC 901
 Qy 996 TTGGAAGAGATGATGATATTTTGAACCAACTGATATTGTGACAGAGAACTTGTGAGAGTG 1055
 Db 902 TTGCGGAGAGAGCAGATGATATTTTGAAGCTTAATCGATCAATTGACAGAAACTGTGCAAGTG 961
 Qy 1056 TCCCTGTTGGCTGTGAGATGAGATGACACAGAGACTGTGCTTGATGAGAGCGACATGTAC 1115
 Db 962 TCCCTGTTGTCTGTGATGGGATGAGATGACCAAGAACTTGTCTCCATGGAACGTGACATGTAC 1021
 Qy 1116 TGAATTAAGCCGAGCCAAGCCCCCTTACAGAGCTTCTTCGCGCCAGTTCCGTGCAAGCC 1175
 Db 1022 TGAAGACGAGCGAGCGCTCAGCGCCCTTACAGAGCTCTTCTGCAAGGTCTGCGCGGCAT 1081
 Qy 1176 TCTTTTATGGGCTCCACCTTCAACATTAAGCTTAAACAAGAGAGATGATTTCAAGCCC 1235
 Db 1082 TCTCTTATGGGCTTCAACCTTCAACATTAAGCTTAAAGAGAGACTTTAAGCTTTGTGCA 1141
 Qy 1236 AATCAGAGAGACGAGAGATGCTACAGCTG-----GCAATCATTTGGCCGCTTC 1283

Dp	1142	AAAGAGAGCGCTGACACGGATGAGAAAGAGTGGCTATTGACGACACCTAAGGCTGCTTC	1201
QY	1284	CTAGCGCTTCAGTCCCATGATCAGATCCTCCGAGGCAAACTCAAGACCAAACTACTG	1343
Dp	1202	TTTGGACTCAACCCAAAATCAACATCTTCTCTTAAAGACTTAAAGCCAAACTATTG	1261
QY	1344	TGCGCCAAAGAGGAATCCCTTCTCCACAGAGGCTCTGCCAAAACCAAGGACCCAA	1403
Dp	1262	TGTTCTTAAAG-----ACCCCTCTCTGAAAGGCAAGTGTAAAGATGCGAACACGAAAAC	1315
QY	1404	CAGAACTTTAGGGGCGAGGAAGACAAACAAAGCTGTGAAGCTTAAAGCTGTGAGCGCTTC	1463
Dp	1316	CAGAAAG-----ATGCTGGAATTTAAAGGCTGGAATTTCTTG	1354
QY	1464	AAGTCTGGGCCCACTGTATCAGAGGCCAGGTAATCAAGTGCCTCCCAACAGCGCCCTAGC	1523
Dp	1355	AAATGTGTTCCAAAGTTTAAAGAGAGGCTCCATTTGGGCCACAGGCACTCAGAGCG	1414
QY	1524	CCCACTCC	1531
Dp	1415	CACCTTAC	1422

Feature	Location/Qualifiers
AA21226	standard; DNA; 16125 BP.
AA21226	
22-NOV-1999	(first entry)
Human CG1CE	genomic DNA sequence.
CG1CE	Best's macular dystrophy; mutation; diagnosis; detection; BMD.
age-related macular dystrophy; ss.	
Homo sapiens.	
Key	
Intron	1..1513
exon	1514..1570
intron	1571..3278
exon	3279..3466
CDS	3467..6616
exon	6617..7227
intron	7228..7461
exon	7462..8353
intron	8354..8508
exon	8509..16125

Query Match	Best Local Similarity	Score	DB 2	Length	16125	Mismatches	Indels	Gaps	0
Sequence 16125 BP, 3988 A; 4175 C; 4215 G; 3726 T; 0 U; 21 Other;	28.8%	642.2	DB 2	16125		3	0	0	
Best Local Similarity	99.5%	Pred. No. 4,7e-139;							
Match 644; Conservative	0;	Mismatches	3;	Indels	0;	Gaps	0;		

QY	1202	CAGCTTGAA	CAAAGAGAGAT	TGAGATTCCAGGCCCAT	CAGAGAGAGAGAGAGAT	TTCTCTA	1261
Db	13916	CAGCTTAA	CAAAAGAGAT	TGAGTTTCCAGGCCCAT	CAGAGAGAGAGAGAT	TTCTCTA	13975
QY	1262	CGCTGGCAT	CAATTGGCGCTT	CTTCCATAGGCT	CGACAGTCCCAT	GATCATCTCTCCAGGGC	1321
Db	13976	CGCTGGCAT	TTTGGCGCTT	CTTAGGCTCGACAGT	CCCATGATCAT	CCATCTCTCCAGGGC	14033
QY	1332	AAACTCAAG	CAACAACTACT	GTGGCCCAAGAGGAAT	CTCTTCTCCA	CAGAGGGCTGGC	1381
Db	14036	AAACTCAAG	CAACAACTACT	GTGGCCCAAGAGGAAT	CTCTTCTCCA	CAGAGGGCTGGC	14092
QY	1382	CAAAACCA	CAGGCGACCC	CAACGAACGTTA	GAGGCGCAGAGAAC	CAACAGGCTGGAA	1441
Db	14096	CAAAACCA	CAGGCGACCC	CAACGAACGTTA	GAGGCGCAGAGAAC	CAACAGGCTGGAA	14152
QY	1442	GCTTAAG	CGTGTGAGCGCTT	CAAGTCTGGCCCA	CGTATCAAGGCGC	AGCTACTACAG	1501
Db	14156	GCTTAAG	CGTGTGAGCGCTT	CAAGTCTGGCCCA	CGTATCAAGGCGC	AGCTACTACAG	14215
QY	1502	TGCCCA	CAGACGCGCT	CAAGCCCACTCCAT	GTTCCTCCCTAGAA	CCATCAGCGC	1561
Db	14216	TGCCCA	CAGACGCGCT	CAAGCCCACTCCAT	GTTCCTCCCTAGAA	CCATCAGCGC	14272
QY	1562	GTCAAG	CTTCA	CAGTGTCA	CAGGCA	TAGACACCAAGCAAAAGCTTAA	GACTGTGAG 1621
Db	14276	GTCAAG	CTTCA	CAGTGTCA	CAGGCA	TAGACACCAAGCAAAAGCTTAA	GACTGTGAG 14333
QY	1622	TTCTGGGG	CCCAAGAAAGTTT	GAAATTGCTCT	CAAGAGAGGATGGGGC	CTTGA	TGGAGCA 1681
Db	14336	TTCTGGGG	CCCAAGAAAGTTT	GAAATTGCTCT	CAAGAGAGGATGGGGC	CTTGA	TGGAGCA 14392
QY	1682	CCCAAG	AGTATCTCA	TGAGAGAGAAAC	TGTGAGATTAA	CCTACGGATATG	CCAGA 1741
Db	14396	CCCAAG	AGTATCTCA	TGAGAGAGAAAC	TGTGAGATTAA	CCTACGGATATG	CCAGA 14455
QY	1742	GATCCCG	AAATCACTCA	AGAACTTTGAA	CAATCCAA	CAATCACTCA	CTAC 1801
Db	14456	GATCCCG	AAATCACTCA	AGAACTTTGAA	CAATCCAA	CAATCACTCA	CTAC 14512
QY	1802	ACTCAAA	GATCACTG	GATCTCTTAT	TGGGCGCTT	GAAACA	GAGGAAG 1848
Db	14516	ACTCAAA	GATCACTG	GATCTCTTAT	TGGGCGCTT	GAAACA	GAGGAAG 14562

RESULT 10

ABAI4559

ID ABAI4559 standard; DNA; 16650 BP.

XX ABAI4559;

XX AC

XX DT 23-JAN-2002 (first entry)

DE Human nervous system related polynucleotide SEQ ID NO 6890.

XX Human; nootropic; neuroprotective; cytosolic; dermatological; virucide;

XX immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;

XX antiparkinsonian; antischlicking; antianemic; antiarthritic; cancer;

XX antitirneumatic; hepatotropic; cerebroprotective; antiinflammatory;

XX antidiabetic; antidiabetic; antileuer; anticonvulsant; antifungal;

XX antiparasitic; cardiant; immune disorder; cardiovascular disorder;

XX neurological disease; infection; nephrotropic; gene therapy; vaccine; db.

OS Homo sapiens.

XX

XX WO200159063-A2.

XX

XX 16-AUG-2001.

XX

XX 17-JAN-2001; 2001MO-US001334.

XX

PR 31-JAN-2000; 2000US-0179065P.

PR	17-NOV-2000;	2000US-0249300P.
PR	01-DEC-2000;	2000US-0250391P.
PR	01-DEC-2000;	2000US-025160P.
PR	05-DEC-2000;	2000US-0251030P.
PR	05-DEC-2000;	2000US-0251988P.
PR	06-DEC-2000;	2000US-0256719P.
PR	06-DEC-2000;	2000US-0251479P.
PR	08-DEC-2000;	2000US-0251856P.
PR	08-DEC-2000;	2000US-0251869P.
PR	08-DEC-2000;	2000US-0251989P.
PR	08-DEC-2000;	2000US-0251990P.
PR	11-DEC-2000;	2000US-0254097P.
PR	05-JAN-2001;	2001US-0259678P.
PA	(HUMA-) HUMAN GENOME SCI INC.	
PI	Rosen CA, Barash SC, Ruben SM;	
PI	WPI; 2001-541565/60.	
XX		
XX		
PT	Nucleic acids encoding 3224 human nervous system antigen polypeptides,	
PT	useful for preventing, diagnosing and/or treating nervous system cancers	
PT	and metastases.	
PS	Disclosure; SEQ ID NO 6888; 1701bp + Sequence Listing; English.	
CC	The invention relates to novel genes (ABAI1004-ABA21534) and proteins	
CC	(ABAI1678-ABAI8001) useful for preventing, treating or ameliorating	
CC	medical conditions e.g. by protein or gene therapy. The genes are	
CC	isolated from a range of human tissues disclosed in the specification.	
CC	The nucleic acids, proteins, antibodies and (ant.)agonists are useful in	
CC	the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and	
CC	ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,	
CC	digestive tract, liver, lung, or urogenital; (b) immune	
CC	diseases e.g. Addison's disease, diabetes mellitus, Crohn's disease,	
CC	anaemia, autoimmune thyroiditis, allergies, autoimmune haemolytic	
CC	multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)	
CC	cardiovascular disorders such as myocardial infarction; (d) wound healing	
CC	; (e) neurological diseases e.g. cerebral palsy and epilepsy; and (f)	
CC	infectious diseases such as viral, bacterial, fungal and parasitic	
CC	infections. Note: The sequence data for this patent did not form part of	
CC	the printed specification, but was obtained in electronic format directly	
CC	from WIPO at ftp.wipo.int/pub/published_pat_sequences	
XX		
SQ	Sequence 18530 BP; 4512 A; 4799 C; 4984 G; 4235 T; 0 U; 0 Other;	
	Query Match 28.8%; Score 642.2; DB 5; Length 18530;	
	Best Local Similarity 99.5%; Pred. No. 4,9e-139;	
	Matches 644; Conservative 0; Mismatches 3; Indels 0; Gaps 0;	
OY	1202 CAGCCTGAACAAAGGAGATGATGAGTTCACAGGCCCAATCAGAGAGCAGAGATGTCTCA	1261
Db	11895 CAGCCTGAACAAAGGAGATGATGAGTTCACAGGCCCAATCAGAGAGCAGAGATGTCTCA	11955
OY	1262 CGCTGGCATCTTTGGCCGCTTCTTAAGGCCTGAGTCCCATGATCACATCTCTCCAGGGG	1321
Db	11955 CGCTGGCATCTTTGGCCGCTTCTTAAGGCCTGAGTCCCATGATCACATCTCTCCAGGGG	12014
OY	1322 AAATCTAAGACCAAATCTACTGTGGCCCAAGAGAGAAATCCCTTCTCAAGAAGGCTTGCC	1381
Db	12015 AAATCTAAGACCAAATCTACTGTGGCCCAAGAGAGAAATCCCTTCTCAAGAAGGCTTGCC	12074
OY	1382 CAAAAAACCAAGGCGACCCAAACAGAACTTAGGGGCGAGAAAGCAACAAGGCTGTGAA	1441
Db	12075 CAAAAAACCAAGGCGACCCAAACAGAACTTAGGGGCGAGAAAGCAACAAGGCTGTGAA	12134
OY	1442 GCTTAAGCTGTGAGCGCTTCAAGTCTGAGGCCCATGTATCAGAGCCAGGCTTACTAG	1501
Db	12135 GCTTAAGCTGTGAGCGCTTCAAGTCTGAGGCCCATGTATCAGAGCCAGGCTTACTAG	12194
OY	1502 TGCCCAACAGAGAGCCCTTATATCCCTATATCCCTATATCCCTATATCCCTATATCC	

Db 12195 TGCCCCACAGAGCCCTCCAGCCCCACCTCCCATGTTCTTCCCTAGAACCATCAGGCGC 12254
Qy 1562 GTCAAGCTTCAGAGTGCACAGGCATAGACACCAAGAAAGCTTAAGACTGTGAG 1621
Db 12255 GTCAAGCTTCAGAGTGCACAGGCATAGACACCAAGAAAGCTTAAGACTGTGAG 12314
Qy 1622 TTCTGGGGCCCAAGAAAGTTTGAATTGCTCTCAGAGAGCATGGGGCTTGATGAGCA 1681
Db 12315 TTCTGGGGCCCAAGAAAGTTTGAATTGCTCTCAGAGAGCATGGGGCTTGATGAGCA 12374
Qy 1682 CCCAAGATATCTCAAGTAGAGAGAGAAACTGTGAGTTTAACCTGACGATATGCCAGA 1741
Db 12375 CCCAAGATATCTCAAGTAGAGAGAGAAACTGTGAGTTTAACCTGACGATATGCCAGA 12434
Qy 1742 GATCCCGGAAATACCTCTCAAGAAACCTTTGGAAACATCAACCAACATACACACTAC 1801
Db 12435 GATCCCGGAAATACCTCTCAAGAAACCTTTGGAAACATCAACCAACATACACACTAC 12494
Qy 1802 ACTCAAGATCACTGATCCTTATTTGGGCTTGAAAAACAGGATG 1848
Db 12495 ACTCAAGATCACTGATCCTTATTTGGGCTTGAAAAACAGGATG 12541

RESULT 12
AB14558
ID AB14558 standard; DNA; 18537 BP.
XX
AC AB14558;
XX
DT 23-JAN-2002 (first entry)
XX
DE Human nervous system related polynucleotide SEQ ID NO 6889.
XX
KW Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnery;
KW antiparkinsonian; anticholinerg; antianaemic; antirheumatic; cancer;
KW antineumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
KW neurological disease; infection; neurotropic; gene therapy; vaccine; ds.
XX
OS Homo sapiens.
XX
PN WO200159063-A2.
PD 16-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001334.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214866P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 23-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234987P.
PR 25-SEP-2000; 2000US-0234988P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0236337P.
PR 29-SEP-2000; 2000US-0236337P.
PR 29-SEP-2000; 2000US-0236357P.
PR 29-SEP-2000; 2000US-0236358P.
PR 29-SEP-2000; 2000US-0236359P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 20-OCT-2000; 2000US-0242211P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.

PR	08-NOV-2000;	2000US-0246527P.
PR	08-NOV-2000;	2000US-0246528P.
PR	08-NOV-2000;	2000US-0246532P.
PR	08-NOV-2000;	2000US-0246609P.
PR	08-NOV-2000;	2000US-0246610P.
PR	08-NOV-2000;	2000US-0246611P.
PR	08-NOV-2000;	2000US-0246613P.
PR	17-NOV-2000;	2000US-0249207P.
PR	17-NOV-2000;	2000US-0249208P.
PR	17-NOV-2000;	2000US-0249209P.
PR	17-NOV-2000;	2000US-0249210P.
PR	17-NOV-2000;	2000US-0249211P.
PR	17-NOV-2000;	2000US-0249212P.
PR	17-NOV-2000;	2000US-0249213P.
PR	17-NOV-2000;	2000US-0249214P.
PR	17-NOV-2000;	2000US-0249215P.
PR	17-NOV-2000;	2000US-0249216P.
PR	17-NOV-2000;	2000US-0249217P.
PR	17-NOV-2000;	2000US-0249218P.
PR	17-NOV-2000;	2000US-0249244P.
PR	17-NOV-2000;	2000US-0249245P.
PR	17-NOV-2000;	2000US-0249264P.
PR	17-NOV-2000;	2000US-0249265P.
PR	17-NOV-2000;	2000US-0249297P.
PR	17-NOV-2000;	2000US-0249299P.
PR	17-NOV-2000;	2000US-0249300P.
PR	01-DEC-2000;	2000US-0250391P.
PR	01-DEC-2000;	2000US-0251160P.
PR	05-DEC-2000;	2000US-0251030P.
PR	05-DEC-2000;	2000US-0251988P.
PR	06-DEC-2000;	2000US-0256719P.
PR	08-DEC-2000;	2000US-0251479P.
PR	08-DEC-2000;	2000US-0251856P.
PR	08-DEC-2000;	2000US-0251869P.
PR	08-DEC-2000;	2000US-0251989P.
PR	08-DEC-2000;	2000US-0251999P.
PR	11-DEC-2000;	2000US-0254097P.
PR	05-JAN-2001;	2001US-0259678P.
PA	(HUMA-) HUMAN GENOME SCI INC.	
XX		
XX	Rosen CA, Barash SC, Ruben SM;	
PI	WPI; 2001-541565/60.	
XX		
DR	Nucleic acids encoding 3224 human nervous system antigen polypeptides,	
PT	useful for preventing, diagnosing and/or treating nervous system cancers	
PT	and metastases.	
PS	Disclosure; SEQ ID NO 6893; 1701bp + Sequence Listing; English.	
XX		
CC	The invention relates to novel genes (ABA11004-ABA21534) and proteins	
CC	(AB11678-AB118001) useful for preventing, treating or ameliorating	
CC	medical conditions e.g. by protein or gene therapy. The genes are	
CC	isolated from a range of human tissues disclosed in the specification.	
CC	The nucleic acids, proteins, antibodies and (ant)agonists are useful in	
CC	the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and	
CC	ovarian cancer and other cancers of the adrenal gland, bone, bone marrow	
CC	disorders e.g. Addison's disease, allergies, autoimmune haemolytic	
CC	anemia, autoimmune rheumatoid arthritis and ulcerative colitis; (c)	
CC	multiple sclerosis, thrombotic thrombocytopenic purpura; (d) wound healing	
CC	cardiovascular disorders such as myocardial ischaemia; (e) cerebral anoxia and epilepsy; and (f)	
CC	infectious diseases such as viral, bacterial, fungal and parasitic	
CC	infections. Note: The sequence data for this patent did not form part of	
CC	the printed specification, but was obtained in electronic format directly	
XX	from Wipo at ftp.wipo.int/pub/published_pct_sequences	
XX		
SQ	Sequence 18537 BP; 4519 A; 4800 C; 4983 G; 4235 T; 0 U; 0 Other;	
Query Match	28.7%; Score 640.6; DB 5; Length 18537;	

	Best Local Similarity	99.4%	Pred. No. 1.2e-138	
Matches	643	Conservative	0	Mismatches 4; Indels 0; Gaps 0;
QY	1202	CAGCTTGAA	CAAAAGAGAGATGAAGTTCCAGCCCAATCAGAGAGACAGAGAGATGCTCA	1261
Db	11901	CAGCTTGAA	CAAAAGAGAGATGAAGTTCCAGCCCAATCAGAGAGACAGAGAGATGCTCA	1196
QY	1262	CGCTGGCATCATTTGGCGCGCTTCCTTAGGCGTGCATGCTCCATATATACCATCTCTCCAGGGC	1321	
Db	11961	CGCTGGCATCATTTGGCGCGCTTCCTTAGGCGTGCATGCTCCATATATACCATCTCTCCAGGGC	1202	
QY	1322	AAACTCAGAGACCAACTACTGTGGCCCAAGAGAGAAATCCCTTCTTCACAGAGGGCCTGCC	1381	
Db	12021	AAACTCAGAGACCAACTACTGTGGCCCAAGAGAGAAATCCCTTCTTCACAGAGGGCCTGCC	1208	
QY	1382	CAAAAAACCA	CAGGACGCCAAACAGAACGTTTAGGGGCCAGAGACCAACAAGGCTTGGAA	1441
Db	12081	CAAAAAACCA	CAGGACGCCAAACAGAACGTTTAGGGGCCAGAGAGACCAACAAGGCTTGGAA	12144
QY	1442	GCTTAAGGCTGTGACACGCTTCAAGTCTGGGCCCATGTATCAGAGGCCATGCTATACAG	1501	
Db	12141	GCTTAAGGCTGTGACACGCTTCAAGTCTGGGCCCATGTATCAGAGGCCATGCTATACAG	12200	
QY	1502	TGCCCCACAGACAGCCCTCTACGCCCACTCCCATGTTCTTCCCTTAGAACCATCAGCGCC	1561	
Db	12201	TGCCCCACAGACAGCCCTCTACGCCCACTCCCATGTTCTTCCCTTAGAACCATCAGCGCC	12260	
QY	1562	GTCAAAGCTTACACAGTGCACAGGCAATAGACACCAAGACCAAAAGCTTAAAGCTGTAG	1621	
Db	12261	GTCAAAGCTTACACAGTGCACAGGCAATAGACACCAAGACCAAAAGCTTAAAGCTGTAG	12320	
QY	1622	TTCTGGGGCCCAAGAAAAGTTTGAATTGCTCTCAGAGAGCGATGGGGCCTTGATGGAGCA	1681	
Db	12321	TTCTGGGGCCCAAGAAAAGTTTGAATTGCTCTCAGAGAGCGATGGGGCCTTGATGGAGCA	12380	
QY	1682	CCCGAAGTATCTCAATATGAGAGAGAGAAAACCTGTGAGCTTTAACCTGACGGATATGCCAGA	1741	
Db	12381	CCCGAAGTATCTCAATATGAGAGAGAGAAAACCTGTGAGCTTTAACCTGACGGATATGCCAGA	12440	
QY	1742	GATCCCGCAAAATCACCTCAAAAGACCTTTGGAAACATACCAACCAACATACACACTAC	1801	
Db	12441	GATCCCGCAAAATCACCTCAAAAGACCTTTGGAAACATACCAACCAACATACACACTAC	12500	
QY	1802	ACTCAAAAGATCACATGAGTCTTATTTGGGCTTGGAAAACAGGGATG	1848	
Db	12501	ACTCAAAAGATCACATGAGTCTTATTTGGGCTTGGAAAACAGGGATG	12547	
RESULT 13				
ABL89697				
ID	ABL89697	standard, cDNA, 1717 BP.		
XX	ABL89697,			
XX	24-MAY-2002	(first entry)		
XX	Human polyomucleotide SEQ ID NO 259.			
XX	Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;			
XX	antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;			
XX	vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;			
XX	cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;			
XX	neurological disease; infection; human; secreted protein; gene; ss.			
OS	Homo sapiens.			
XX	WO200190304-A2.			
XX	29-NOV-2001.			
XX	18-MAY-2001, 2001WO-US016450.			
XX	19-MAY-2000, 2000US-0205515P.			

XX (HUMA-) HUMAN GENOME SCI INC.
 PA Birse CE, Rosen CA;
 PI WPI; 2002-122018/16.
 DR P-PSDB; ABB89288.
 XX Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
 PT prevention of neural, immune system, muscular, reproductive,
 PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
 PT disorders.
 XX Claim 4; SEQ ID NO 259; 2081bp + Sequence Listing; English.
 PS
 XX The invention relates to novel genes (ABL89449-ABL90853) and proteins
 CC (ABB89040-ABB90444) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
 CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
 CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
 CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
 CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
 CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
 CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
 CC cardiovascular disorders such as myocardial ischaemia; (d) wound healing
 CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
 CC infectious diseases such as viral, bacterial, fungal and parasitic
 CC infections. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 1717 BP; 464 A; 460 C; 417 G; 367 T; 0 U; 9 Other;

Query Match 26.4%; Score 587.8; DB 6; Length 1717;
 Best Local Similarity 98.8%; Pred. No. 9.8e-127;
 Matches 600; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 1242 GAGACGAGGAGGATGCTCAGCGCTGGCATTTGGCCGCTTCTAGGCTGCAGTCCCAT 1301
 DB 1 GAGACGAGGAGGATGCTCAGCGCTGGCATTTGGCCGCTTCTAGGCTGCAGTCCCAT 60
 QY 1302 GATCACCATCTCTCCAGGGCAAACTCAAGACCAAACTAGTGTGGCCCAAGAGGAATCC 1361
 DB 61 GATCACCATCTCTCCAGGGCAAACTCAAGACCAAACTAGTGTGGCCCAAGAGGAATCC 120
 QY 1362 CTCTCCACGAGGGCTGCCCCAAAACCAAGGACGCAAAAGAGGTTAGGGGCGCAG 1421
 DB 121 CTCTCCACGAGGGCTGCCCCAAAACCAAGGACGCAAAAGAGGTTAGGGGCGCAG 179
 QY 1422 GAAGACAAAGAGGCTGGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGA 1481
 DB 180 GAAGACAAAGAGGCTGGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGA 239
 QY 1482 CAGAGGCGAGGCTCTACAGTCCCAAGAGGCGCCCTCAGGCCCACTCCCATGTTCTC 1541
 DB 240 CAGAGGCGAGGCTCTACAGTCCCAAGAGGCGCCCTCAGGCCCACTCCCATGTTCTC 299
 QY 1542 CCCCTAGAACCATCAGGCGCCCTCAAGGCTTCAAGTGTCAAGGATAGACCAACCAAGAC 1601
 DB 300 CCCCTAGAACCATCAGGCGCCCTCAAGGCTTCAAGTGTCAAGGATAGACCAACCAAGAC 359
 QY 1602 AAAAGCTTAAAGACTGTGAGTTCGTGGGGCCCAAGAAAGTTTGAATTTCTTCAAGAGC 1661
 DB 360 AAAAGCTTAAAGACTGTGAGTTCGTGGGGCCCAAGAAAGTTTGAATTTCTTCAAGAGC 419
 QY 1662 GATGGGGCTTATGATGAGACCCAGAGATCTCAAGTGAAGAGAAAGCTGTGAGATTT 1721
 DB 420 GATGGGGCTTATGATGAGACCCAGAGATCTCAAGTGAAGAGAAAGCTGTGAGATTT 479
 QY 1722 AACCTGACGATATGACAGAGATCCCGAAATATCACTCAAGAAAGCTTTTGAACAATCA 1781

DB 480 AACCTGACGATATGACAGAGATCCCGAAATATCACTCAAGAAAGCTTTGGAACAATCA 539
 QY 1782 CCACCAACATACACACTACACTCAAGATCATGATCTTATTTGGGCTTGGAAAC 1841
 DB 540 CCACCAACATACACACTACACTCAAGATCATGATCTTATTTGGGCTTGGAAAC 599
 QY 1842 AGGGATG 1848
 DB 600 AGGTCTG 606

RESULT 14
 AB061177
 ID AB061177 standard; cDNA; 2112 BP.
 XX
 AC AB061177;
 XX
 DT 26-FEB-2003 (first entry)
 XX
 DE FLJ20132 fls clone COL06441 encoding sequence.

XX Neuroprotective; immunomodulator; cancer; chromosome 19; cytostatic;
 KW anti-inflammatory; gene therapy; nutritional supplement; wound; burn;
 KW ulcer; Alzheimer's disease; Huntington's disease;
 KW amyotrophic lateral sclerosis; autoimmune disorder; inflammation;
 KW vulnery; gene; ss.

XX Homo sapiens.
 XX MO200231111-A2.
 XX
 PD 18-APR-2002.
 XX
 PP 11-OCT-2001; 2001MO-US027760.
 XX
 PR 12-OCT-2000; 2000US-00687527.
 XX
 PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
 PI Xue Ao, Yang Y, Wehrman T, Drmanac RT;
 DR WPI; 2002-426278/45.
 DR N-PSDB; ABP43933.
 XX

PT New polypeptides and their encoded proteins, useful as nutritional
 PT sources or supplements, or in gene therapy, particularly for treating
 PT wounds, Alzheimer's disease, amyotrophic lateral sclerosis, cancer or
 PT inflammation.
 XX
 PS Claim 1; SEQ ID # 390; 357bp + Sequence Listing; English.

XX The invention relates to 446 newly isolated polynucleotide sequences. The
 CC activity of polynucleotides of the invention may be described as,
 CC inflammatory, neuroprotective, immunomodulator, cytostatic and anti-
 CC inflammatory. Compositions comprising nucleic acids of the invention are
 CC useful for treating a mammalian subject, or as nutritional sources or
 CC supplements. These are useful in gene therapy, particularly for treating
 CC wounds, burns or ulcers, Alzheimer's disease, Huntington's disease,
 CC amyotrophic lateral sclerosis, autoimmune disorders, cancer or
 CC inflammation. The nucleic acids and polypeptides are also useful in
 CC diagnostic and research methods. The sequences given in records AB060788-
 CC AB061233 represent polynucleotides of the invention. NOTE: The sequence
 CC data for this patent did not form part of the printed specification, but
 CC was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX

Query Match 25.9%; Score 577.2; DB 6; Length 2112;
 Best Local Similarity 69.9%; Pred. No. 3.2e-124;
 Matches 793; Conservative 0; Mismatches 338; Indels 3; Gaps 1;

QY 103 CCATGACATCACTTACACAAAGCAAGTGTATATGCCGCTTAGGCTCTTCCTCCGCC 162
 DB 190 CGATGACCGCTCACTTACACAGCCGAGTGGCGAAGCCGCCCTTCGATGGCTTCCAGC 249
 QY 163 TGTGCTGTGCTGGCGGGGACAGATCTTACAGCTGTATATGCGAGTTCTTAATCTTCC 222
 DB 250 TGTGCTGTGCTGGCGGGGACAGATCTTACAGCTGTATATGCGAGTTCTTAATCTTCC 309
 QY 223 TGTGCTGTGCTGTAT 282
 DB 310 TTGGGTTCTACATGGCGGTGATGTGTGCTTACCGCTTGTGTGTGTGTGTGTGTGTGT 369
 QY 283 TGATGTTGAGAAATGATCTGTATGTCGACAGCTATACATCAGCTATCCCATTTTCTT 342
 DB 370 GCTACTTGAGAAAGCTTGTGATTTATTTGTGACAGTATGTCAGCTATCCCTGTCTCT 429
 QY 343 TGTGCTGTGCTGTAT 402
 DB 430 TGTGCTGTGCTGTAT 489
 QY 403 TGTGCTGTGCTGTAT 462
 DB 490 TGTGCTGTGCTGTAT 549
 QY 463 AAGCCGCGCTGTGCGGGGACAGCTATCCGCTTACCGCAACCTGGGCAACGCTCTATCC 522
 DB 550 GCGGCCCGCTTACCGGGGACAGCTATGCTGTACAGCGAGGCTCTCGCCGCTCTATCC 609
 QY 523 TGTGCTGTGCTGTAT 582
 DB 610 TGTGCTGTGCTGTAT 669
 QY 583 CAGGCTTTATGATCTCGGCGAGAACACAAAGCTGTGAGAACTGTAGCCACACAAACA 642
 DB 670 CTGGGTTTATGACCCCGAGAGACCGCAAAAGTTGAAACCTGAACCTATCTTACAAACA 729
 QY 643 TGTGCTGTGCTGTAT 702
 DB 730 AGTACTGTGCTGTAT 789
 QY 703 GAATCGGGGACCTTATCTGTCTCAAGAGCTGTGTGAAGATGAAACCTTGTGCTATC 762
 DB 790 GCAATCGGCGAACAGCGCCCTTAAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 849
 QY 763 AGTGTGACACCTGTATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 822
 DB 850 AATGTGTGATGTCTTTTACATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 909
 QY 823 TGACTGT 882
 DB 910 TGACATGT 969
 QY 883 CAGCCAAAGCTTACCTGT 942
 DB 970 CGGCTCAGGGTTTAAAGAACCAACCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1029
 QY 943 AGT 1002
 DB 1030 AGT 1089
 QY 1003 AGATGT 1062
 DB 1090 AGGACGATGT 1149
 QY 1063 TGTGCTGT 1122
 DB 1150 TGTGCTGT 1209
 QY 1123 AGCCCGAGGACAGCCCGCTTACAGCTGTCTTCCGCTGTGTGTGTGTGTGTGTGTGTGT 1179
 DB 1210 CAGCCGAGGCTGT 1269
 QY 1180 TTATGT 1233

DB 1270 TCAGGGCTTCCACTTTGACATCAGCTGTGGCCAAAGAGATCATGATTCAGC 1323
 RESULT 15
 ID AB280973 standard; DNA; 2137 BP.
 AC AB280973;
 DT 15-OCT-2003 (first entry)
 DE Human bestrophin (vitelliform macular dystrophin 2-like protein 1) gene.
 KW Bestrophin; vitelliform macular dystrophin 2; VMD2; ds; gene; vaccine;
 KW antioestrogen; antidiabetic; immunomodulator; hypotensive; cardiac;
 KW antileptic; osteopathic; antineoplastic; cytostatic; obesity;
 KW energy homeostasis; metabolism; triglyceride; body-weight regulation;
 KW eating disorder; cachexia; diabetes mellitus; hypertension; gallstones;
 KW coronary heart disease; hypercholesterolemia; osteoarthritis; cancer;
 KW sleep apnea; chromosome 19.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 212..1741
 FT /*tag= a
 FT /product= "VMD2-like protein 1"
 WO2003030922-A2.
 PD 17-APR-2003.
 PF 09-OCT-2002; 2002MO-BE011321.
 PR 09-OCT-2001; 2001BP-00124059.
 (DEVE-) DEVELOPMENTAL BIOLOGICAL FORSCH.
 PI Steuermangel A, Broenner G, Fritsch R, Eulenberger K, Ciosek T;
 DR WPI; 2003-393411/37.
 DR P-PSDB; ABR58056.
 PS New pharmaceutical compositions comprising a Bestrophin gene, polypeptide
 PT or nucleic acid, for treating, alleviating and/or preventing metabolic
 PT diseases, e.g. obesity, cachexia, diabetes mellitus, hypertension, or
 PT gallstones.
 XX Claim 2; Fig 4c; 85pp; English.
 XX This sequence represents the coding region for a member of the human
 CC bestrophin gene family designated vitelliform macular dystrophin 2 (VMD2)
 CC -like protein 1. The dystrophin gene family are involved in energy
 CC homeostasis and metabolism of triglycerides. The sequence can be used for
 CC the manufacture of an agent for detecting and/or verifying, for treating,
 CC alleviating and/or preventing disorders including metabolic diseases such
 CC as obesity and other body-weight regulation and related disorders such as
 CC eating disorder, cachexia, diabetes mellitus, hypertension, coronary
 CC heart disease, hypercholesterolemia, osteoarthritis, gallstones, cancers
 CC of the reproductive organs, and sleep apnea. The gene is found on human
 CC chromosome 19
 SQ Sequence 2137 BP; 399 A; 671 C; 622 G; 445 T; 0 U; 0 Other;
 Query Match 25.9%; Score 577.2; DB 8; Length 2137;
 Best Local Similarity 69.9%; Pred. No. 3.2e-124;
 Matches 793; Conservative 0; Mismatches 338; Indels 3; Gaps 1;
 QY 103 CCATGACATCACTTACACAAAGCAAGTGTATATGCCGCTTAGGCTCTTCCTCCGCC 162
 DB 210 CGATGACCGCTCACTTACACAGCCGAGTGGCGAAGCCGCCCTTCGATGGCTTCCAGC 269

Search completed: March 25, 2005, 23:51:32
Job time : 1172.53 secs

Qy	163	TGCTGCTGCTGCTGGGGGAGAGATCTTAAAGAGCTGCTATATGGGAGATTTCTTATCTTCC	222
Dp	270	TGCTGCTACTGTGGGTGGAGCACTTCAAACTTCGTGTGGCGAGAGGTGCTGTGCTTCC	329
Qy	223	TGCTCTGTACTACATCATCCGCTTTATTTATTAAGGCTGGCCCTCAACGAAACAAGC	282
Dp	330	TTGGGTTTACAATGAGCGCTGAGTGTGCTGCTTACCGCTTTGTGTGTACGAAAGGAGAAC	389
Qy	283	TGATGTTTGAAGAACTGACTCTGTATTTGCGACAGCTACATCCAGCTCATTCCTTCT	342
Dp	390	GCTACTTGAGAGAGTTGTGATTTATTTATGTGACAGTATGCAAGCTCATTCCTGTCTCT	449
Qy	343	TGTGTCTGGGCTTCTACGTGACGTGTGTCTGTGACCCGCTGGTGGAAACAATGACGAAC	402
Dp	450	TCGTGCTTGGCTTTTATGTGACGCTGTGTGTGAACCGCTGTGTGGCAATTACTTATGCA	509
Qy	403	TGCGGTGGCCCGACCGGCTCATGTAGCCGTGTGTGGGGCTTGCTGTGAAGCAAGACGAGC	462
Dp	510	TGCGGCTGCCGAGCGGCTCATGTGCGTGTGTGGGGGACCGTGTGACGAGCGACGACCC	569
Qy	463	AAGCGCGGCTCTGTGGGCGGACGCTCATCCGCTTACGCCAACCTGGGCAACCTGTCTATCC	522
Dp	570	GGCGCGGCTCTACCGGGGACACATCATGTGCTTACGACGAGGGCTCTCGGCGGTGTCTATCC	629
Qy	523	TGCGAGGCTGAGCACCGCACTTACAAGCCTTCCAGGCGCCACGACCTGTGTCAAG	582
Dp	630	TGCGGCTCGGTGAGCACCGCGGGTGTCAAGCGCTTCCACATATGACCACTGTGTGAGG	689
Qy	583	CAGGCTTATGATCTCCGCGACAAACAAGCAGTTGTGAGAAACTGAGCTTACCAACAACA	642
Dp	690	CTGGGTTTATATACCGCGACGAGGCGCAAGAAAGTTGAAACCTGAACTCATCTTACACA	749
Qy	643	TGTTCTGGGGTCCCGCTGGGTGTGTGTTTGTGCAACTGTTCATATGAAGCGTGTGTGAGATC	702
Dp	750	AGTACTGGGTGCTCCGCTGTGTGTTCTTCAACTGTGGGCAACAGCCCTACGCGAGGGCC	809
Qy	703	GAATCCGGGACCTTATCCTGTCTCCAGAGCCTGTCTAAAGAGATGAAACACTTGTGCTATC	762
Dp	810	GCATCCGGGACAAACGCGCCCTTAAAGCTGTGTCTGAGAGAGCTGAATGTTTTGGGGCA	869
Qy	763	AGTGTGACACTGTATGCTTACATGCACTGTGATTTATATCCACTGTGTATATACAGGTGG	822
Dp	870	AATGTGATGCTCTTTCACATATGACTGAAATTAAGGTACCCCTGTGTATACAGCAGGTGG	929
Qy	823	TGACTGTGGCGGTGATACAGCTTCTTCCGTGACTGTGTCTATGGGGCGGAGTTTCTGAACC	882
Dp	930	TGACCATGCACTGTACAGCTACTTCTGTGGCTGTCTATTTGTGTGCCAGTTTCTTGAACC	989
Qy	883	CAGCCAAAGGCTTACCCTGGGCATGAGCTGTGACCTGTGTGCGCGCTTCAAGTCTCTGC	942
Dp	990	CGGCTCAGGGTTACAAABAACACAGACTTAAAGCTGTGTGTGTCCTTCAACCTCTTGGC	1049
Qy	943	AGTTCTTCTTGTATGTTGGCTGTGTGAGAGGTGGCAGACAGCTCATCAACCCCTTTGGAG	1002
Dp	1050	AGTTCTTCTTCTACGCGCGGCTGTGCTCAAGGTATGCTGAGCACTCATCAACCCCTTGGAG	1109
Qy	1003	AGGATGATGATGATTTTGAAGCAACATGGAATTTGTGAGAGAAATTTGAGAGTGTCCCTGT	1062
Dp	1110	AGGACGATGATGATCTTTGAGACCAACTTTCATGTGATGATGAACCTTCCAGGTGTCCATGC	1169
Qy	1063	TGGCTGTGGATGAGATGACACGAGACTGTCCGATGGAGCCGGAATGTACTGGAATA	1122
Dp	1170	TGGCATGTGACGAGATGTATGTATGACTGTGTGTGTGTGTGAAAGGACTTTGTACTGTGATG	1229
Qy	1123	AGCCGAGCCACAAGCCCCCTTACACAGCTGTCTT---GCCCCAGTTCCGTGTGAGCTCTT	1179
Dp	1230	CAGCCGAGGCTGTGCGCCCATATACACAGGCGCTATCTTCAAGCTGTGGGAGCCTTCTT	1289
Qy	1180	TTATGTGGCTTCACTTCAACATCAAGCTGTGAACAAAGAGATGAGATTTCCAGC	1233
Dp	1290	TTCAAGGGCTCACTTTTGAATCAAGCTGTGGCCAAAGAAAGATGAGATTTCCAGC	1343

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 26, 2005, 07:39:17 ; Search time 1241.65 Seconds
(without alignments)
10698.156 Million cell updates/sec

Title: US-09-622-964A-2

Perfect score: 2229

Sequence: 1 caggagagctccaccagccta.....aaaaaaaaaaaaaaaa

Scoring table: IDENTITY_NUC

Searched: 5552208 seqs, 2979665951 residues

Total number of hits satisfying chosen parameters: 11104416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 19: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
- 21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*
- 22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1754.8	78.7	1758	US-10-492-032-1	Sequence 1, Appl1
2	1247.2	56.0	1263	US-09-746-783-3	Sequence 3, Appl1
3	587.8	26.4	1717	US-10-264-237-859	Sequence 259, App
4	577.2	25.9	2137	US-10-492-032-3	Sequence 3, Appl1
5	576.8	25.9	1530	US-10-492-032-30	Sequence 129, Ap
6	459.2	20.6	2500	US-10-108-2604-1299	Sequence 129, Ap
7	446.2	20.0	1422	US-10-492-032-7	Sequence 7, Appl1
8	423	19.0	2028	US-10-492-032-5	Sequence 5, Appl1
9	350.4	15.7	1326	US-10-071-766-10	Sequence 10, Appl
10	215.4	9.7	1350	US-09-768-826-16	Sequence 16, Appl
11	215.4	9.7	1350	US-10-874-484-16	Sequence 16, Appl

C 12	213	9.6	620	10	US-09-814-353-19504	Sequence 19504, A
C 13	189.2	8.5	1292	14	US-10-198-846-11070	Sequence 11070, A
C 14	175	7.9	1198	9	US-09-880-107-2174	Sequence 2174, Ap
C 15	161.4	7.2	578	13	US-10-027-632-134530	Sequence 134530, A
C 16	161.4	7.2	578	13	US-10-027-632-134530	Sequence 134530, A
C 17	149.4	6.7	748	13	US-10-027-632-134529	Sequence 134529, A
C 18	149.4	6.7	748	13	US-10-027-632-134529	Sequence 134529, A
C 19	135.2	6.1	539	16	US-10-029-386-7583	Sequence 7583, Ap
C 20	134.2	6.0	235	16	US-10-029-386-21283	Sequence 21283, A
C 21	125.4	5.6	2914	18	US-10-723-860-7441	Sequence 7441, Ap
C 22	125.4	5.6	2914	18	US-10-723-860-8303	Sequence 8303, Ap
C 23	116	5.2	853	14	US-10-198-846-11346	Sequence 11346, A
C 24	115.4	5.2	1954	14	US-10-723-860-7576	Sequence 7576, Ap
C 25	104	4.7	462	14	US-10-198-846-2561	Sequence 2561, Ap
C 26	103.6	4.6	3186778	13	US-10-027-632-134561	Sequence 134561, A
C 27	103.6	4.6	3186778	17	US-10-027-632-174961	Sequence 174961, A
C 28	102	4.6	615	14	US-10-198-846-7843	Sequence 7843, Ap
C 29	98	4.4	507	10	US-09-764-872-195	Sequence 195, App
C 30	74.2	3.3	239	18	US-10-637-855-110	Sequence 110, App
C 31	66.2	3.0	442	10	US-09-814-353-14343	Sequence 14343, A
C 32	66	3.0	748	13	US-10-027-632-134529	Sequence 134529, A
C 33	66	3.0	748	17	US-10-027-632-134529	Sequence 134529, A
C 34	65.4	2.9	260	10	US-09-814-353-1601	Sequence 1601, Ap
C 35	65.4	2.9	260	10	US-09-814-353-13563	Sequence 13563, A
C 36	65.4	2.9	463	10	US-09-814-353-13563	Sequence 13563, A
C 37	64	2.9	751	14	US-10-198-846-1862	Sequence 1862, Ap
C 38	63.2	2.8	298	10	US-09-814-353-806	Sequence 806, App
C 39	63.2	2.8	298	10	US-09-814-353-7178	Sequence 7178, Ap
C 40	60	2.7	60	10	US-09-908-975-16358	Sequence 16358, A
C 41	54.4	2.4	65	10	US-09-908-975-29941	Sequence 29941, A
C 42	52.4	2.4	481	18	US-10-767-701-19477	Sequence 19477, A
C 43	49.6	2.2	1274	18	US-10-437-963-87249	Sequence 87249, A
C 44	49.2	2.2	1263	10	US-09-746-783-3	Sequence 3, Appl1
C 45	49.2	2.2	1326	13	US-10-071-766-10	Sequence 10, Appl1

ALIGNMENTS

RESULT 1
US-10-492-032-1
Sequence 1, Application US/10492032
Publication No. US20050049212A1
GENERAL INFORMATION:
APPLICANT: Steuermagel, Arnd
APPLICANT: Brommer, Gunter
APPLICANT: Fritsch, Rudiger
APPLICANT: Eulenberg, Karsten
TITLE OF INVENTION: Closek, Thomas
TITLE OF INVENTION: Restriction and Restriction Homologous Proteins Involved in the
FILE REFERENCE: 2923-606
CURRENT APPLICATION NUMBER: US/10/492,032
CURRENT FILING DATE: 2004-04-08
PRIOR APPLICATION NUMBER: PCT/EP02/11321
PRIOR FILING DATE: 2002-10-09
PRIOR APPLICATION NUMBER: EP01124059.5
NUMBER OF SEQ ID NOS: 37
SOFTWARE: PatentIn version 3.2
SEQ ID NO 1
LENGTH: 1758
TYPE: DNA
ORGANISM: Homo sapiens
US-10-492-032-1
Query Match 78.7%; Score 1754.8; DB 19; Length 1758;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1756; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Oy 105 ATGACATCACTTACACAGCAAGTGGCTATATCCGCTTATGCTCTTCTCCGCTG 164
Db 1 ATGACATCACTTACACAGCAAGTGGCTATATCCGCTTATGCTCTTCTCCGCTG 60

165 CTGCTGTGGGCGGGGAGCATCTCAAGCTGTATGCGAGTTCTTAATCTTCCTG 224
61 CTGCTGTGGGCGGGGAGCATCTCAAGCTGTATGCGAGTTCTTAATCTTCCTG 120
225 CTCTGTACTACATCATCCGCTTTATTTATAGGCTGGCCCTCAGGAAACAAGCTG 284
121 CTCTGTACTACATCATCCGCTTTATTTATAGGCTGGCCCTCAGGAAACAAGCTG 180
285 ATGTTTGAAGAACTGACTGTATTTGCGACAGCTACATCCAGTATCCCAATTCCTC 344
181 ATGTTTGAAGAACTGACTGTATTTGCGACAGCTACATCCAGTATCCCAATTCCTC 240
345 GTGCTGGGCTTTACGTGAGCGTGTGCTGACCCGCTGGTGGAGCAAGTACAGAACTG 404
241 GTGCTGGGCTTTACGTGAGCGTGTGCTGACCCGCTGGTGGAGCAAGTACAGAACTG 300
405 CCGTGGCCCGACCGGCTCTATGAGCGTGTGCGGCTTGTGGAAGGCAAGGACGAGCA 464
301 CCGTGGCCCGACCGGCTCTATGAGCGTGTGCGGCTTGTGGAAGGCAAGGACGAGCA 360
465 GGCCTGGCTGTGCGGCGGACGCTCATCCGCTACGCAACCTGGGCAAGTGTCTCTG 524
361 GGCCTGGCTGTGCGGCGGACGCTCATCCGCTACGCAACCTGGGCAAGTGTCTCTG 420
525 CGCAGCGTTCAGCAGCGGAGTCTACAGGCGCTTCCCGACGCGCAGCAGCTGGTGAACA 584
421 CGCAGCGTTCAGCAGCGGAGTCTACAGGCGCTTCCCGACGCGCAGCAGCTGGTGAACA 480
585 GGCCTTTATGACTCCCGGAGAAACAAGAGTTGGAGAAACTGAGCTTACACACAACTG 644
481 GGCCTTTATGACTCCCGGAGAAACAAGAGTTGGAGAAACTGAGCTTACACACAACTG 540
645 TTCTGGGTCCTCGGCTGTGTTGCTCAACCTGTCAATGAAGCGTGGCTTGAAGTGA 704
541 TTCTGGGTCCTCGGCTGTGTTGCTCAACCTGTCAATGAAGCGTGGCTTGAAGTGA 600
705 ATCCGGGACCTTATCTGCTCAAGGCTGTGGAAGGAGATGAACACCTTGGCTGACTG 764
601 ATCCGGGACCTTATCTGCTCAAGGCTGTGGAAGGAGATGAACACCTTGGCTGACTG 660
765 TGTGGACACTGTATGCTTACAGCTGATTTAGTATCCCACTGTGTATACACAGTGTG 824
661 TGTGGACACTGTATGCTTACAGCTGATTTAGTATCCCACTGTGTATACACAGTGTG 720
825 ACTGTGGCGGTGTACAGCTTCTCTGACTGTATGTTGGGCGGAGTTTCTGAACCA 884
721 ACTGTGGCGGTGTACAGCTTCTCTGACTGTATGTTGGGCGGAGTTTCTGAACCA 780
885 GCCAAGGCTTACCTGGCCATGAGCTGGAACCTGTGTGCGGCTTCTCAAGTTCTGAG 944
781 GCCAAGGCTTACCTGGCCATGAGCTGGAACCTGTGTGCGGCTTCTCAAGTTCTGAG 840
945 TTCTTCTTCTATGTTGGCTGTGCTGAAGTGGCAAGAGCTCATCAACCTTTGGAGAG 1004
841 TTCTTCTTCTATGTTGGCTGTGCTGAAGTGGCAAGAGCTCATCAACCTTTGGAGAG 900
1005 GATGATGATGATTTTGAAGACCACTGATTTGTGCAAGAGATTTTGAAGTGTCCCTGTG 1064
901 GATGATGATGATTTTGAAGACCACTGATTTGTGCAAGAGATTTTGAAGTGTCCCTGTG 960
1065 GCTGTGATGATGATGCAACAGAGCTGTGCTGAGTGAAGCGGACATGTAAGTGAATAG 1124
961 GCTGTGATGATGATGCAACAGAGCTGTGCTGAGTGAAGCGGACATGTAAGTGAATAG 1020
1125 CCGAGGCAAGAGCCCTCTACAGAGCTGTTCCGCGCAGTTCCGTGAGGCTCTTATAG 1184
1021 CCGAGGCAAGAGCCCTCTACAGAGCTGTTCCGCGCAGTTCCGTGAGGCTCTTATAG 1080
1185 GGTCTCACTTCAACATCAAGCTGAACAAAGAGAGATGAGTTTCAAGCCCAATCAAGAG 1244
1081 GGTCTCACTTCAACATCAAGCTGAACAAAGAGAGATGAGTTTCAAGCCCAATCAAGAG 1140

1245 GACGAGAGAGATGCTCAAGCTGGCATTTGGCCGCTTCTTACGCGCTGAGTCTCCATGAT 1304
1141 GACGAGAGAGATGCTCAAGCTGGCATTTGGCCGCTTCTTACGCGCTGAGTCTCCATGAT 1200
1305 CACCATCTCTCCAGGAGGCAAACTCAAGAGACCAATCTAAGTGTGCGCAAGAGGAAATCCCTT 1364
1201 CACCATCTCTCCAGGAGGCAAACTCAAGAGACCAATCTAAGTGTGCGCAAGAGGAAATCCCTT 1260
1365 CTCCAGAGAGGCTTGGCCCAAAACCAAGAGCGCCCAACAGAAAGTGTAGGCGCCAGGAA 1424
1261 CTCCAGAGAGGCTTGGCCCAAAACCAAGAGCGCCCAACAGAAAGTGTAGGCGCCAGGAA 1320
1425 GACAACAGAGGCTTGAAGCTTGAAGGCTGTGAGCGCTTCAAGTCTGGCCCACTGTATCAG 1484
1321 GACAACAGAGGCTTGAAGCTTGAAGGCTGTGAGCGCTTCAAGTCTGGCCCACTGTATCAG 1380
1485 AGGCGAGCTACTACAGTGTGCGCCCAAGAGCGCCCTCAGCCCACTCCATGTTCTTCCCC 1544
1381 AGGCGAGCTACTACAGTGTGCGCCCAAGAGCGCCCTCAGCCCACTCCATGTTCTTCCCC 1440
1545 CTGAACCATCAAGCGCGCTGAAAGCTTCAAGTGTCAAGGCTTGAACACCAAGACAA 1604
1441 CTGAACCATCAAGCGCGCTGAAAGCTTCAAGTGTCAAGGCTTGAACACCAAGACAA 1500
1605 AGCTTAAAGACTGTGAGTCTTGGGCGCAAGAAAGTTTGAATTTGCTCTCAGAGAGCAT 1664
1501 AGCTTAAAGACTGTGAGTCTTGGGCGCAAGAAAGTTTGAATTTGCTCTCAGAGAGCAT 1560
1665 GGGGCTTGTATGAGACACCCAGAAAGTATCTCAAGTGAAGAGAAACTGTGAGTTTAAAC 1724
1561 GGGGCTTGTATGAGACACCCAGAAAGTATCTCAAGTGAAGAGAAACTGTGAGTTTAAAC 1620
1725 CTGACGATATGACAGAGATCCCGAAATCACTCAAGAAACCTTTGGAACATACCA 1784
1621 CTGACGATATGACAGAGATCCCGAAATCACTCAAGAAACCTTTGGAACATACCA 1680
1785 ACCAATACACACTACACTCAAGATCAATGATCTTATTTGGGCTTGGAAAAACAG 1844
1681 ACCAATACACACTACACTCAAGATCAATGATCTTATTTGGGCTTGGAAAAACAG 1740
1845 GATGAAGACATTTCTTAA 1862
1741 GATGAAGACATTTCTTAA 1758

RESULT 2
US-09-746-783-3
Sequence 3, Application US/09746783
Publication No. US20030044935A1
GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
McCoy, John M.
Lavalley, Edward R.
Racie, Lisa A.
Trecey, Maurice
Spaulding, Vikki
Agostino, Michael J.
Howes, Steven H.
Fechtel, Kim
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
ENCODING THEM
NUMBER OF SEQUENCES: 231
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: MA
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

Query Match	25.9%	Score 577.2;	DB 19;	length 2137;
Best Local Similarity	69.9%;	Pred. No. 2e-154;		
Matches 793; Conservative	0;	Mismatches 338;	Indels 3;	Gaps 1;

QY	103	CAATGACCACTACATTAAACAAAGCAAGTGGCTAATGCCCCGCTTAAGCTCTTCTCCGCC	167
Db	210	CGATGACCGTGCACCTTACACAGCCGAATGGCGAAACGCCCTTCGGAGGCTTCCCGACG	267
QY	163	TGCTGCTGTCTGGCGGGGAGCATCTACAAGCTGCATATAGCGAGTTCCTTAATCTTCC	222
Db	270	TGCTGTACTGTGGCGGTGGGAGGATCTACAAATCTCTGTGGCAGAGCTCTCTGCTTCC	328
QY	223	TGCTGTGCTACTACATCATCCGCTTTATATTAAGCTGGCCCTCACGGAAGAACAAAGC	282
Db	330	TTGGGGTTCTACATAGCGCGCTGAGTGGTCCCTTACCGCTTGTGCTGACCCGAAGGCGAGAGC	389
QY	283	TGATGTTTGGAGAACTGACTCTGTATTTGGCGACGCTACATCCAGCTCATCCCATTTTCC	342
Db	390	GCCTACTTCAGAAAGCTTGTGATTTATTTGTGACCAAGATGCAAGCTTCATCCCTGTCTCT	449
QY	343	TGCTGTGAGGCTTTCTACGTTGACGCTGTCGTGACCCGCTGTGTGAAACAGTACGAAACC	402
Db	450	TGCTGTGCTTGGCTTTTATGTGACGCTGTGTGTGAACCGCTGTGTGAGCCAGTACTTAATGA	509
QY	403	TGCGCGTGGCCGACCGCCCTCATGAGCGCTGGTGTTCGGGCTTTCGTGAAAGCAGAGAGAGC	462
Db	510	TGCCGCTGCCCGAAGCGCTCATGTGTGTGTGTGGCGGGCACCGTGCACGAACGGAGAAC	569
QY	463	AAAGCCGAGCTGTGTGCGGCGCAAGCTCATCCGCTACGCGCAACCTGGGCAACGTTCTATCC	522
Db	570	GGCGCGCGCTTACCGGGCGCACTCATGCGTATGACAGGGGCTTCGGCGGTGTCTATCC	629
QY	523	TGCGGAGCGCTCAGCACCCGAGTCTACAAAGCGCTTCCCGACGCGCCAGCACTTGTGTCAAG	582
Db	630	TGCGCTCGGTGACACCGCGGTGTTCAAAGCGCTTCCCAACATAGACCACTGTGTGTGAGG	689
QY	583	CAGGCTTATGACTCCGGGAGAAACAAGCAATGTGAGAAATGAGGCTTACACACACAA	642
Db	690	CTGGGTTTATGACCCGCGAGGAGCGCAAGAAAGTTGAAAACCTGAACTTCATCTTCAACA	749
QY	643	TGTTCTGGGCTGCCCTGGGCTGTGGTTTGCCAACTGTCAATGAAAGCGTGGCTTGGAGGTC	702
Db	750	AGTACTGGGCTGCCCCGTGTGTTCTCCAACCTGGCGGCAACAGCCCGAAGCGAGAGGCC	809
QY	703	GAATCCGGGAGCCCTATCTCTGCTCCAGAGCCCTGTGAAAGAAATGAAACCTTGGCTACTC	762
Db	810	GCATCCGGGACAAACGCCCCCTTAAAGCTGCTGTGAGAGGACTGAAATGTTTTCCGGGCA	869
QY	763	AGTGTGACACCTGATAGCTTACAGACTGATGATTAAGTATCCCACTGGTGTATACACAGGTGG	822
Db	870	AATGTGGAATGCTCTTTCACATATGACTGGAATTAGAGTAAACCCCTGTGTACACGACGAGTGG	929
QY	823	TGACTGTGGCGGTGTACAGCTTCTTCTCTGACTTGTGTATGATTGGGCGGAGTTTCTGAAAC	882
Db	930	TGACCATGTCACTGTACAGCTACTTCCCTGGCTTCCCTCATATTGTGTGCGCAGTTCCGAGAC	989
QY	883	CAGCCAAAGGCTTACCTGGGACATAGCGTGAACCTGTGTGGCCCGTCTTACAGTTCTCTGC	942
Db	990	CGGCTAGAGGTTAAACAAAGACCAACGACCTTACAGCTGTGTGTGTGCACATCTTCAACCTCTTGG	1049

Qy	943	AGTCTTCTTCATGTTGGCTGGAGTGGAGACCACTATCAACCCCTTGGAG	1002
Db	1050	AGTCTTCTTCTTAGCGCGGCTGGCTCAAGAGTACCTGACACTATCAACCCCTTGGAG	11099
Qy	1003	AGGATGATGATGATTTTGAGACCAACTGGATTGTGACAGAGAAATTGCA93GTGTCCTGT	10622
Db	1110	AGGACGATGATGACTTTGAGACCAACTTCTGATGATAGAGAACTTCCAG3GTGTCATGTC	11669
Qy	1063	TGGCGTGTGATGATGATGTCACAGGACCTGCGCTCGGATGTGAGACCGGACATATGACTGGATA	11222
Db	1170	TGGCAGTGTGACGAGATGTATGATGACTCTGGCTGTCTGGAGAGAGACTTGTGACTTGGGATG	12299
Qy	1123	AGCCCGAGCCACAGCCCTCTTACACAGCTGTCTT---CGCCCAGTTCCTGTGAGACTTCT	11799
Db	1230	CAGCCGAGGCTCTCGGCCCCCATACACAGCGGCTACTGTCTTCCAGCTGGG3CAGGCTTCTCT	12899
Qy	1180	TTATGGGCTCCACCTTCAACATCAGCTCGATCAAGAGAGAGATGAGTTCAGC	1233
Db	1280	TTCAAGGCTCCACCTTTGACATCAGCTCTGGCCCAAGAGAGATGACAGTTCAGC	1343

RESULT 5

```

: Sequence 30 Application US/10492032
: Publication No. US20050049212A1
: GENERAL INFORMATION:
: APPLICANT: Steuernagel, Arnd
: APPLICANT: Bronner, Gunter
: APPLICANT: Fritsch, Rüdiger
: APPLICANT: Eulenberg, Karsten
: APPLICANT: Closssek, Thomas
: TITLE OF INVENTION: Beetrophin and Beetrophin Homologous
: TITLE OF INVENTION: Regulation of Energy Homeostasis
: FILE REFERENCE: 2923-606
: CURRENT APPLICATION NUMBER: US/10/492,032
: CURRENT FILING DATE: 2004-04-08
: PRIOR APPLICATION NUMBER: PCT/EP02/11321
: PRIOR FILING DATE: 2002-10-09
: PRIOR APPLICATION NUMBER: EP01124059.5
: PRIOR FILING DATE: 2001-10-09
: NUMBER OF SEQ ID NOS: 37
: SOFTWARE: PatentIn version 3.2
: SEQ ID NO 30
: LENGTH: 1530
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-10-492-032-30

```

Query Match	25.9%	Score 576.8;	DB 19;	Length 1530;
Best Local Similarity	70.0%	Pred. No. 2.1e-154;		
Matches 792; Conservative	0;	Mismatches 337;	Indels 3;	Gaps 1;

QY	105	ATGACCATCACTTATACAAAGCCAAAGGGGTATAGCCCGCTTAGGCTCTTTCGCGCTG	164
Db	1	ATGACCGTACCTTACACAGCCCGAGTGGGAAGCCCGCTTGGTGGCTTTCACAGCTG	60
QY	165	CTGCTGTCTGGCGGGGCGAGCATCTACAAGCTGCTATATAGCGAGTTCTTAAATCTTCG	224
Db	61	CTGCTACTGTGGCGGGGAGCATCTACAAGCTCTGTGGCGAGAGCTGCTCTTCCTT	120
QY	225	CTCTGCTCTATCATCATCCGCTTTTATTTATPAGGCGCGCCTCACGGGAAGAACAGCGTG	284
Db	121	GGTTCTTACATGCGCGCTAGTGTCTCTACCGCTTTGTGCTGACCGAAGGCGAAGCCG	180
QY	285	ATGTTTGAGAACTGACTCTGTATTTGCGACAGCTACATCAGCTCATGCCCATTTTCCTTC	344
Db	181	TACTTCGAGAACTGTGATTTATTTGTGACAGATGACAGAGCTCATCCCTGTCTCTTC	240
QY	345	GGGCTGGGCTTTAGCTGACGCGTGGGCGAGACCCGCTGGTGGAAACCAATACGAAACCTG	404
Db	241	GTCCTGGGCTTTTATGTGACGCGTGGTGGTGAACCGCTGTGGAGCCAGTATCTATGCTG	300
QY	405	CCGTGCGCCGACCGGCTCATGAGCTGTGTGTGGGCTTGTGTGAGAGCAAGACGAGCA	464

Dp	301	CCGCTGCCCGACGGGCTCATGTGGGTGGCGGGACCGTGTGCACAGACGGACGACGCCG	360
Qy	465	GGCCGGCTGTGCGGCGCACGCTCATCCGTACGCAACCTGGGCAACGTGCTCATCTGT	524
Dp	361	GGCGCGCTCTACCGGCGCACACTATGTGGCTACGAGAGGGCTCTCGGCGGTGTCTACTGT	420
Qy	525	CGACAGCTCAGACACCGCACTTACAGGGCTTCCCGACGGCGCCACACTGTGTGCAGCA	584
Dp	421	CGCTCCGTCAGCACCGGGGTGTTCAAGGGCTTCCCGACATAGACACAGTGGTGAAGCT	480
Qy	585	GGCTTTATGATCCGGCGCAACACAGAGATTGGAGAACTGAGCCTTCCACACAACATGT	644
Dp	481	GGGTTTATGACCCCGGAGAGCGCAAGAGTTTGAAATCTGAACTCACTCTACCAACAG	540
Qy	645	TTCTGGGTGCCCCGTGGTGTGTTGGCAACCTGTCAATGAAGCGGTGGCTTGAAGTGA	704
Dp	541	TACTGGGTGGCCCTGCGTGTGGTTCTCCACCTGGCGGACAGGGCCGACGGAGGGCGCG	600
Qy	705	ATCCGGGACCTTATCTGCTTCAGAGCCTGTGTAAACGATGAACACTTGGGTACTCAG	764
Dp	601	ATCCGCGCAACACACGCGCTTAACTGTGTCTCAGAGCGCTGAATGTTTTTCGGGGCAA	660
Qy	765	TGTGACACCTGTATGCTTACATGGAATTATCCACGTGTATACAGGTTGATG	824
Dp	661	TGTGAAATGCTCTTTCATCTATGACTGTGAATGACGTAACCTCTGTGTACACGAGTGTG	720
Qy	825	ACTGTGCGGGTGTACACCTTCTCTGTACTGTATGTTGGGCGGCAATTTCTGAACCA	884
Dp	721	ACATGTGCACGTGTACAGTACTTCTGTGGCTTGCTCAATTGGTCCGCAATTCGTGAACCG	780
Qy	885	GCCAAAGCCTTACCCTGGGCATGAGCTGAGCCTGTGTGTGCCGTTCACGTTCTCTGAG	944
Dp	781	GCTCAGAGTTTACAAGAACAACGACCTTGAACCTGTGTGGCCATCTTCAACCTCTTGGAG	840
Qy	945	TTCTTTTCTTATGTGGGTGGCTGAAGTGGCAGAAGCTTCAACACCCCTTGGGAAG	1004
Dp	841	TTCTTCTTCTAAGCCGCGGTGGCTCAAGGTACTAGACGCTCATCAACCCCTTGGGAAG	900
Qy	1005	GATATATGATGATTTTGTAGAACCACTGATTTGTGACAGGAATTTGTGAGGTGTCCCTGTTG	1064
Dp	901	GACGATGATGACTTTGAGACCACTTTCTGTATCATATAGAAATCTTCAGGTGTCTCATGTCTG	960
Qy	1065	GCTGTGATGATGATGACACAGAGACTGTGCTTGATGAGACCGGACATGTACTGTGAATAAG	1124
Dp	961	GCAGTGTGACGAGATGTATGATGACTGTGGCTGTGTGGAAGAGCACTTGTACTGGGATGCA	1020
Qy	1125	CCGAGGCCACAGCCCCCTTACACAGCTGTTC---CGCCAGTTTCCGTGAGCTCTTCTT	1184
Dp	1021	GCCAGGCTGCGCGCCCATACACAGCGGCTTACTGTCTTCAGCTGCGGACCTTCTTTC	1080
Qy	1182	ATGGGCTTCACCTTCAACATCAGGCTGTGAACAAAGAGGATGAGATTCCAG	1233
Dp	1081	CAGGGCTTCACCTTTGACATACCGCTGGCCAAAGAGACATGCAATTTCCAG	1132

RESULT 6

```

: Sequence 1299, Application US/10108260A
: Publication No. US20040005560A1
: GENERAL INFORMATION:
: APPLICANT: HELIX RESEARCH INSTITUTE
: TITLE OF INVENTION: NO. US20040005560A1e1 full length cDNA
: FILE REFERENCE: H1-A0106
: CURRENT APPLICATION NUMBER: US/10/108,260A
: CURRENT FILING DATE: 2002-03-27
: NUMBER OF SEQ ID NOS: 5458
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 1299
: LENGTH: 2500
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-10-108-260A-1299

```


Query Match 20.6%; Score 459.2; DB 17; Length 2500;
 Best Local Similarity 63.5%; Pred. No. 1,7e-120;
 Matches 701; Conservative 0; Mismatches 403; Indels 0; Gaps 0;

```

QY 105 ATGACCATCTTACACAGCCAGTGGCTAATGCCGCTTAGGCTCTCTCCGCTG 164
DB 229 ATGACTGTCACTTACCTCAAGTAAAGTAGCAATCTTTTGGATTTCATAGTTA 288
QY 165 CTGCTGTGCTGGGGGAGCATCTCAAGCTCTATATGCGAGTTCTTATCTTCCTG 224
DB 289 CTCTCAAGTGGAGGAGGAGCATCTCAAACTACTAGCAAGGAAATTTATGTTTGTCT 348
QY 225 CTCTGCTACTACATCATCCGCTTATTTATAGGCTGAGCTCCCTCAAGGAAACAAGCTG 284
DB 349 GTTCTTTATACAGCAATTAAGTTTGTGTACAGATTCTTATCAGAGATCCAAAAGCT 408
QY 285 ATGTTGAGAAATGACTCTGTATTCGACAGCTATACCACTCACTCCATTCCTTCCTC 344
DB 409 TACTTGAATAAATTAATCAATTAAGTGTGACAGATATGCGAACTAATTCAGTAACTTT 468
QY 345 GTGCTGGGCTTCTACGTGACGCTGTGTCGACCCGCTGGTGAACAGTACAGAACTG 404
DB 469 GTGCTTGGGTTTATATTTACTGTGATGAAACCATGTGGAGACAGTTGTGATTTG 528
QY 405 CCGTGGCCGAGCCGCTCATGAGCTGTGTGGGCTTCTGCGAAGGAGAGAGAGCA 464
DB 529 CCGTGGCCGAGAGAGGCTAATGTTCCATCTCAAGAGTTCACGAAAGAGAGAGAC 588
QY 465 GGCAGGCTCTGCGGGGAGAGCTCATCCGCTTACGCAACTGGGCAAGCTGTCTCTG 524
DB 589 GGGCGGCTCTTGAAGAGAGCTGATGCGCTAGCGCAATCTCACTCCCTGCTCATCTT 648
QY 525 CCGAGCGTCAAGCCGAGCTTCAAGGCTTCCCGAGGCGCCAGCACTGTGTGAAGCA 584
DB 649 CGCTCGGTGAGCACTGTGTGACAAAGATTTCCCAATAGAGCACTAGTTGAAGCA 708
QY 585 GGCCTTATGACCTCGGAGAGAGACAGAGAGTTGGAACTGAGCTTACACACACATG 644
DB 709 GGTTTATGACAAAGATGAAGAAATTTATCAACCACTCAAGTCTCTCATCTGAAA 768
QY 645 TTCTGGGTCCTGAGGTGTGTTGCCAAGCTGTCAATGAGAGCGTGGCTTGAAGTGA 704
DB 769 TATGAGTTCCATTCATCTGTTTGGAAATCTTGGAATCAAGCCGAGATGAAGTGA 828
QY 705 ATCCGAGACCTTATCTCTCTCAAGAGCTGTGAACGAGATGAACACCTTGGTACTG 764
DB 829 ATCAAGACAGTGTGATCTGCAATCATTTGATGACTGAATGAATGATACCGCTTGG 888
QY 765 TGTGACACCTGTATGCTACAGACTGATTAAGTATCCCACTGTGTATACAGAGTGTG 824
DB 889 TGCAGCTCTTATTCGTTATGACTGGGTTGGAAATCCGCTGTTTACACCCAGTGTG 948
QY 825 ACTGTGGGTGTA CAGTTCTTCTGACTGTCTAGTTGGGCGGAGTTTGAAGCA 884
DB 949 ACTGTGTGCTATACCTTCTTCTTGGCGCTGATTTGAAGCCAGTTTGTGATCC 1008
QY 885 GCGAAGGCTTACCTGAGAGAGTGAAGTGAAGAGCTCATCAACCCCTTGAAGAG 944
DB 1009 ACCAAAGCTTACGAGGAGAGTGTGATCTTTTCAATCCATCTTCAACCTCTTCAA 1068
QY 945 TTCTTCTTATGTTGGCTGAGAGTGAAGTGAAGAGCTCATCAACCCCTTGAAGAG 1004
DB 1069 TTCTTCTTATGAGAGATGGCTTAAAGTGAAGAGAGCTTATCAACCTTGTGAAGAA 1128
QY 1005 GATGATATGATTTTGAAGCACTGATGTGTGACAGAAATTTGAGGTGCTCTGTTG 1064
DB 1129 GATGATATGATTTTGAAGCACTGATGTGTGACAGAAATTTGAGGTGCTCTGTTT 1188
QY 1065 GCTGTGATGAGATGACAGAGAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1124
DB 1189 GCTGTGAGAGATGACAGAGAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1248

```

```

QY 1125 CCGAGCCAGAGCCCTTACAGAGCTGTTCGCCCCAGTTCCGTGAGGCTCTTATG 1184
DB 1249 TCTGTGTGTCGCCCAACATACATGAGTGGAGCTGTACTAGTACATACCTCATTTCTG 1308
QY 1185 GAGTCCACCTTCAACATCAGCTG 1208
DB 1309 GGGTCAACAGTCCAGATGGGCTG 1332

```

RESULT 7

```

US-10-492-032-7
; Sequence 7, Application US/10492032
; Publication No. US20050049212A1
; GENERAL INFORMATION:
; APPLICANT: Steuernagel, Arnd
; APPLICANT: Bronner, Gunter
; APPLICANT: Fritsch, Rudiger
; APPLICANT: Eulenberg, Karsten
; APPLICANT: Ciossek, Thomas
; TITLE OF INVENTION: Bestrophin and Bestrophin Homologous Proteins Involved in the
; TITLE OF INVENTION: Regulation of Energy Homeostasis
; FILE REFERENCE: 2923-606
; CURRENT FILING DATE: 2004-04-08
; PRIOR APPLICATION NUMBER: PCT/EP02/11321
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: EP01124059.5
; PRIOR FILING DATE: 2001-10-09
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7
; LENGTH: 1422
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-492-032-7

```

Query Match 20.0%; Score 446.2; DB 19; Length 1422;
 Best Local Similarity 63.8%; Pred. No. 6.7e-117;
 Matches 751; Conservative 0; Mismatches 378; Indels 48; Gaps 3;

```

QY 105 ATGACCATCTTACACAGCCAGTGGCTAATGCCGCTTAGGCTCTCTCCGCTG 164
DB 1 ATGACGTTTCAATACCTCTCAAGTGGCGAGGCCCTTCGAGTTCTCTGCGCTG 60
QY 165 CTGCTGTGCTGGGGGAGAGCTTACAAAGCTGTATATGAGAGTCTTATCTTCTG 224
DB 61 CTCTTCGCTGAGAGGAGAGATCTACAGCTCTCTTCAAGAAATTTCTCTTGGG 120
QY 225 CTCTGCTACTACATCATCCGCTTATTTATAGGCTGAGCTTCAAGAGAAACAAGCTG 284
DB 121 GCGTGTACGCTGTGCTTATGATCACTACCGGCTGTGCTGACCCAGAGACAGATAC 180
QY 285 ATGTTGAGAAATGACCTGTATTTGCGACAGCTATCCAGCTCATCTCCATTTCTTC 344
DB 181 GTGTATGCTCAGGTGCGGAGTGAACCGCTCAAGAGCTCATATCTCTTGTCTTT 240
QY 345 GTGCTGGGCTTCTAGTGAAGCTGTGCTGACCCGCTGTGGAACCAAGTGAAGAGAG 404
DB 241 GTATTTGGGTTCTATGTAATCTGTGTGTGAACCGCTGTGTGTCTCCAGTACAAAGATC 300
QY 405 CCGTGGCCGAGCCGCTTATATAGCTGTGTGAGGCTTCTGAAAGGCAAGAGAGCA 464
DB 301 CCGTGGCCGAGCCGCTTATATAGCTGTGTGAGGCTTCTGAAAGGCAAGAGAGCA 360
QY 465 GGCAGGCTGTGCGGGGAGAGCTCATCGTGAAGGCAACCTGGGCAACAGTGTCTATCTG 524
DB 361 GGCAGGCTGTGCGGGGAGAGCTCATCGTGAAGGCAACCTGGGCAACAGTGTCTATCTG 420
QY 525 CCGAGGCTGAGCCAGAGTCTTACAGAGCTTCCCAAGGCGCCAGAGCACTGTGTGAAGCA 584
DB 421 CGCTCGGTGAGCAACCGGCTGTGATGAGGCTTCCCAAGGCGCCAGAGCACTGTGTGAAGCA 480
QY 585 GGCCTTATGACTCGGAGAGAAACAGAGAGTGTGAAGAACTGAGCTTACCAACAACTG 644

```

Db 481 GGTTCATGTCCAGAGAGAGAGAAAAGTTTGAGACCTGAATCCGACTTCAACAAG 540
Qy 645 TTTCGGGCGCCCTGGGTGTGTTGGCCAACTGTCAATGAAGCGGTGTGAGGTGCA 704
Db 541 TACTGGGTCCTCGTGTGTGTCAACCACTGGGCGCCGAGAGGAGCGGGGA 600
Qy 705 ATCCGGAGCCCTATCTGCTCCAGAGCCTGTGTAAGAGATGAACACTTGTGACTAG 764
Db 601 ATACGTACAGATATGCTCTGTCTACTTTTGGAGAGCTGAACAGTATCCAGCCAG 660
Qy 765 TGTGACACCTGTATGCTTACAGATGATATGATATCCCACTGTGTATACACAGGTGTG 824
Db 661 TGCAGACATGATATCCATATGATGATGATGATGATGATGATGATGATGATGATG 720
Qy 825 ACTGTGGGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 884
Db 721 ACCATAGCGGTCTACTCTTCTTGTGCGCTCTGCTGCTGCTGCTGCTGCTGCTG 780
Qy 885 G-----CCAGGCGCTTACCCTGGCC 903
Db 781 GAGCAGAGGAGCTGCCAACTCAAGAGCTTGAAGCCAGGCGCAGAGCCAGCCAGCC 840
Qy 904 AT----GAGCTGAACCTGTGTGCGCGCTTCAAGTTCCTGCAAGTTCCTTCTATGT 959
Db 841 CTGGAGAGCCCGACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
Qy 960 GCGTGGCTGAAGGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1019
Db 901 GCGTGGCTGAAGGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 960
Qy 1020 GAGACCAACTGATTTGTGACAGAGATTTGACAGAGTGTGCTGTGCTGTGATGATG 1079
Db 961 GAGACCAACTGATTTGTGACAGAGATTTGACAGAGTGTGCTGTGATGATGATG 1020
Qy 1080 CACCAGAGCTGCTCGGATGAGAGCCGACATGTATGATGATGATGATGATGATGATG 1139
Db 1021 TACCAGAACTTCCCGCGCTGAGAGAGCCAGATGATGATGATGATGATGATGATG 1080
Qy 1140 CCGTACAC--AGCTGCTCGCGCCAGTTCGCTGAGGCTCTTATGATGATGATGATG 1196
Db 1081 CCGTACAC--AGCTGCTCGCGCCAGTTCGCTGAGGCTCTTATGATGATGATGATG 1140
Qy 1197 AACATCAGCCTGAACAAAGAGATGATGATGATGATGATGATGATGATGATGATG 1233
Db 1141 AACCTGGCATGAG 1177

RESULT 8
US-10-492-032-5
Sequence 5, Application US/10492032
Publication No. US20050049212A1
GENERAL INFORMATION:
APPLICANT: Steuermagel, Arnd
APPLICANT: Bronner, Gunter
APPLICANT: Fritsch, Rüdiger
APPLICANT: Eulenberger, Karsten
APPLICANT: Closser, Thomas
TITLE OF INVENTION: Bestrophin and Bestrophin Homologous Proteins Involved in the
FILE REFERENCE: 2923-606
CURRENT APPLICATION NUMBER: US/10/492,032
CURRENT FILING DATE: 2004-04-08
PRIOR APPLICATION NUMBER: PCT/EP02/11321
PRIOR FILING DATE: 2002-10-09
PRIOR APPLICATION NUMBER: EP01124059.5
NUMBER OF SEQ ID NOS: 37
SOFTWARE: PatentIn version 3.2
SEQ ID NO 5
LENGTH: 2028
TYPE: DNA
ORGANISM: Homo sapiens

US-10-492-032-5
Query Match 19.0%; Score 423; DB 19; Length 2028;
Best Local Similarity 62.2%; Pred. No. 3,7e-110;
Matches 697; Conservative 0; Mismatches 400; Indels 24; Gaps 1;
Qy 105 ATGACATCACTTACACAGCAAGAGATGATGATGATGATGATGATGATGATGATGATG 164
Db 1 ATGACATCACTTACACAGCAAGAGATGATGATGATGATGATGATGATGATGATGATG 60
Qy 165 CTGCTGCTGCGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 224
Db 61 CTGCTGCTGCGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
Qy 225 CTGCTGCTGCGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 284
Db 121 GTTCTTATACAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
Qy 285 ATGTTGAGAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 344
Db 181 TACTTGAATAATATCAATTTACTGACAGATATGCTGAACAAATTCAGTAACTTT 240
Qy 345 GTGCTGAGCTTTTACGATGATGATGATGATGATGATGATGATGATGATGATGATG 404
Db 241 GTGCTGAGCTTTTACGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
Qy 405 CCGTGGCGGACCGCTTACAGCTGATGATGATGATGATGATGATGATGATGATGATG 464
Db 301 CCGTGGCGGACCGCTTACAGCTGATGATGATGATGATGATGATGATGATGATGATG 360
Qy 465 GAGCGGCTGCTGCGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 524
Db 361 GAGCGGCTGCTGCGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
Qy 525 CCGAGCTGACAGCAGCTTACAGAGCTTCCCGAGCGGAGAGAGAGAGAGAGAGAGAG 584
Db 421 CCGTGGCTGAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
Qy 585 GCGTGTATGATCTCGGAGAGAACCAAGATGATGATGATGATGATGATGATGATGATG 644
Db 481 GGTATATGACAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
Qy 645 TTTCGGTGCCTGGGTGTGTTGGCCAACTGTCAATGAAGCGGTGTGAGGTGCA 704
Db 541 TATTGGTTCATTTACTGTGTTGGAAATCTTGCAATGAAGCGGAAATGAAGTGA 600
Qy 705 ATCCGGAGCCCTATCTGCTCCAGAGCCTGTGTAAGAGATGAACACTTGTGACTAG 764
Db 601 ATCAGAGACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
Qy 765 TGTGACACCTGTATGCTTACAGATGATGATGATGATGATGATGATGATGATGATG 824
Db 661 TGCAGCTCTTATTCGGTTATGATGATGATGATGATGATGATGATGATGATGATG 720
Qy 825 ACTGTGGGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 884
Db 721 ACTTGTGCTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
Qy 885 GCCAAGGCTTACCTGTGACATGATGATGATGATGATGATGATGATGATGATGATG 944
Db 781 ACCAAGGCTTACCTGTGACATGATGATGATGATGATGATGATGATGATGATGATG 840
Qy 945 TTCTTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 980
Db 841 TTCTTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900
Qy 981 CAGCTATCAACCCCTTGGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1040
Db 901 CAGCTATCAACCCCTTGGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 960
Qy 1041 AGGAATTTGAGAGTGTGCTGTGCTGTGATGATGATGATGATGATGATGATGATGAT 1100
Db 961 AGGAATTTGAGAGTGTGCTGTGCTGTGATGATGATGATGATGATGATGATGATGAT 1020

QY 1101 GAGCCGACATGTAAGTAAGCCGAGCCAGCCCTTACAGAGCTGCTCCGCC 1160
DB 1021 AAGAGGACATTTACTGAGCACTTCTGCTGCTGCGCCACCAATGAGCACTGCTGCT 1080
QY 1161 CATTCCGCTGAGCCTCTTTATGAGCTCCACCTTCAACAT 1201
DB 1081 GACTACTGATACCTTCATCTTCTGAGGCTCAACAGTCCAGAT 1121

RESULT 9

US-10-071-766-10/c
; Sequence 10, Application US/10071766
; Publication No. US20020192678A1
; GENERAL INFORMATION:
; APPLICANT: Huel-Mei Chen
; TITLE OF INVENTION: GENES EXPRESSED IN SENESENCE
; FILE REFERENCE: PA-0043 US
; CURRENT APPLICATION NUMBER: US/10/071,766
; NUMBER OF SEQ ID NOS: 144
; SOFTWARE: PERL Program
; SEQ ID NO 10
; LENGTH: 1326
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020192678A1 1138151.2
US-10-071-766-10

Query Match 15.7%; Score 350.4; DB 13; Length 1326;
Best Local Similarity 99.7%; Pred. No. 2.1e-89;
Matches 351; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1850 AGCATTCTTAACTCTTCTTAAATGAGGATGCTTGGCCAGCAGTCTCAGCTGTGT 1909
DB 1326 AGCATTCTTAACTCTTCTTAAATGAGGATGCTTGGCCAGCAGTCTCAGCTGTGT 1267
QY 1910 GTACACGAGCAGCAGCAGTATGCTTAAATGAGGATGCTTGGCCAGCAGTCTCAG 1969
DB 1266 GTACACGAGCAGCAGCAGTATGCTTAAATGAGGATGCTTGGCCAGCAGTCTCAG 1207
QY 1970 TCTTAAACAGCCTGATTAATGAGGATGCTTAAATGAGGATGCTTGGCCAGCAGTCT 2029
DB 1206 TCTTAAACAGCCTGATTAATGAGGATGCTTAAATGAGGATGCTTGGCCAGCAGTCT 1147
QY 2030 CCTTAAATGCTTAAATGAGGATGCTTAAATGAGGATGCTTGGCCAGCAGTCTT 2089
DB 1146 CCTTAAATGCTTAAATGAGGATGCTTAAATGAGGATGCTTGGCCAGCAGTCTT 1087
QY 2090 AACTGAGCTGATTAATGAGGATGCTTAAATGAGGATGCTTGGCCAGCAGTCTT 2149
DB 1086 AACTGAGCTGATTAATGAGGATGCTTAAATGAGGATGCTTGGCCAGCAGTCTT 1027
QY 2150 CACTTAAATGCTTAAATGAGGATGCTTAAATGAGGATGCTTGGCCAGCAGTCTT 2201
DB 1026 CACTTAAATGCTTAAATGAGGATGCTTAAATGAGGATGCTTGGCCAGCAGTCTT 975

RESULT 10

US-09-768-826-16
; Sequence 16, Application US/09768826
; Patent No. US20020012966A1
; GENERAL INFORMATION:
; APPLICANT: Shi et al.
; TITLE OF INVENTION: 18 human secreted proteins
; FILE REFERENCE: PFS12P1
; CURRENT APPLICATION NUMBER: US/09/768,826
; CURRENT FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: PCT/US00/22350
; PRIOR FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: 60/148,759

; PRIOR FILING DATE: 1999-08-16
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 16
; LENGTH: 1350
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1135)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1148)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1166)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1174)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1181)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1209)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1229)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1266)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1290)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1295)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1305)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1324)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1339)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1341)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1343)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1345)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1348)
; OTHER INFORMATION: n equals a,t,g, or c
; OTHER INFORMATION: n equals a,t,g, or c
US-09-768-826-16

Query Match 9.7%; Score 215.4; DB 9; Length 1350;
Best Local Similarity 61.9%; Pred. No. 1.3e-50;
Matches 420; Conservative 0; Mismatches 211; Indels 48; Gaps 3;

QY 603 GAAACAGCACTTGAAGAACTGAGCTTACCAACAACATGTTCTGAGTCCCTGAGT 662
DB 1021 AAGAGGACATTTACTGAGCACTTCTGCTGCTGCGCCACCAATGAGCACTGCTGCT 1080
QY 1161 CATTCCGCTGAGCCTCTTTATGAGCTCCACCTTCAACAT 1201
DB 1081 GACTACTGATACCTTCATCTTCTGAGGCTCAACAGTCCAGAT 1121

Db 23 GGAGCGGTGGGGTTTGAGAGCGCTGAAATCCGACTTCAACAAAGTACTGGGTCCTCGCGTC 82
Qy 663 TGGTTTGGCAACCTGTCAATGAAAGCGGTGGAGGTGCAATCCGGGACCTATATCTG 722
Db 83 TGGTTTCAACCACTGGCGCGCCCAAGGCCGGAGGAGCGGGCAATACGTAAGATATGCT 142
Qy 723 CTTCAGAGCCCTGTGAACGAGATGAACACTTGGCTACTGCTGTGGAACACTGTATGCC 782
Db 143 CTCTGTCTACTTTTGAAGAGCTGAACAGTACGAGCAAGTGCAGATGCTATTTCCAC 202
Qy 783 TAGCACTGATTAATATCCCACTGCTATATACAGAGTGTGATCTGTGGCTGTACAGC 842
Db 203 TATGACTGATCAGACATCCCTCGTCTACACCAAGTGAGCAATAGCCGTACTCT 262
Qy 843 TTCTTCTGACTTGTCTAGTTGGGCGGCACTTCTGAACCCAG----- 885
Db 263 TTCTTTGGCCCTCTCCCTGGTTGGCGCGCAATTTGTGAGGCCAGAGGCAAGGCGGCTGCCAA 322
Qy 886 -----CGAAGGCTTACCTGCGCATGAG---CTGGAAGCTC 917
Db 323 CCTCAGAAAGTTCTGAAGCCAGGCGAGGACCGCCAGCCCTGGGAGACCCGGAATG 382
Qy 918 GTTGTGCCCCCTTTCAGCTTCTGTGAGTCTTCTTCTATATGTTGGCTGTGAGTGGCA 977
Db 383 TACGTGCTCTCACCACCTCTGCTCAGTCTTCTTCTATGCTGCTGCTCAGAGTGGCT 442
Qy 978 GAGCAGCTCATCAACCCCTTTGAGAGAGATGATGATTTTGAAGCAACTGATTTGTC 1037
Db 443 GAACAGATCATCAACCACTTGTGAGATGATGACCTTTGAGCAAACTACCTCATTA 502
Qy 1038 GACAGAAATTTGAGAGTGTCCCTGTGGCTGTGATGATGACCAAGACCTGCTGG 1097
Db 503 GACCGCAACTTGCAGAGTGTCTCTGTATCGGTGACGAAATGTACAGAACCTTCCGCC 562
Qy 1098 ATGAGCCCGAGCATGTACTGGAATTAAGCCGACCAAGCCCTCTAC---AGTGTCT 1154
Db 563 GCTGAGAGGAGCACTGAGATGAGGACCGGACCCCTTACNCTGTGGCCACG 622
Qy 1155 TCGGCCAGTTCGTTGAGAGCTCCCTTATGAGGCTCAACCTTCAACATCAGCTGAACAA 1214
Db 623 GCGGCCAGATCTTGGCGCCCTCATCTCTGGGCTCACCTTCAACCTGCGCATGAGGAC 682
Qy 1215 GAGGAGATGAGATTCAGC 1233
Db 683 GACCCCTAGCAGAGCCCTGC 701

RESULT 11
US-10-874-484-16
; Sequence 16, Application US/10874484
; Publication No. US20040235113A1
; GENERAL INFORMATION:
; APPLICANT: Shi et al.
; TITLE OF INVENTION: 18 human secreted proteins
; FILE REFERENCE: PFS12P1
; CURRENT APPLICATION NUMBER: US/10/874,484
; CURRENT FILING DATE: 2004-06-24
; PRIOR APPLICATION NUMBER: US/09/768,826
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: PCT/US00/22350
; PRIOR FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: 60/148,759
; PRIOR FILING DATE: 1999-08-16
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 1350
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1135)
; OTHER INFORMATION: n equals a,t,g, or c

; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1148)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1166)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1174)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1181)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1209)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1229)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1266)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1285)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1287)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1290)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1295)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1305)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1324)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1339)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1341)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1343)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1345)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1348)
; OTHER INFORMATION: n equals a,t,g, or c
; US-10-874-484-16

NUMBER OF SEQ ID NOS: 14084
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11070
LENGTH: 1292
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: 1287..1288, 1289, 1290, 1291, 1292
OTHER INFORMATION: n = A,T,C or G
US-10-198-846-11070

Query Match 8.5%; Score 189.2; DB 14; Length 1292;
Best Local Similarity 62.4%; Pred. No. 4.2e-43;
Matches 296; Conservative 0; Mismatches 178; Indels 0; Gaps 0;

Qy 349 TGGGCTTCTACGTGACGCTGTGTGACCCCGTGTGAACCAAGTACAGAACTGCGCT 408
Db 318 TAGGGTTTATGTACTCTGTAGTGAACCGATGGTGAACCAAGTTTGTGAATTTGCCCT 377
Qy 409 GGGCCGACCGGCTCATAGAGCTGTGTGGCTTGTGAGAGGCAAGGACGAGGCGC 468
Db 378 GGCACAGACGCTATATGTTCTCATCTAGCATGTTCACGGAAGCAGACGACGCGGC 437
Qy 469 GCGTCTGCGGCGCAGCTCATTCGCTACGCCAACCTGGCAACGTCCTCATCTCGCGCA 528
Db 438 GCGTCTTATGAAGACGCTGATGCGCTACGTCATCACTCCCTGCTCATCTTTGCGCT 497
Qy 529 GCGTACGACCGCAGTCTTACAGCGCTTCCCGAGCGCCAGCAGCTGTGTGAAGAGGCT 588
Db 498 CGGTAGACACTGTGTGTATCAAAAGATTCCCAATGAGACGATGCTGAAGCAGGTT 557
Qy 589 TTAATGACCTCCGCGACAGACAGAGCTGTGAGAACTGAGCTTACCAACCAATGTTCT 648
Db 558 TTAATGACAGAGATGAAGAAATTTATTCACCACTCAAGTCTCTCATCTGAATATTT 617
Qy 649 GGGTGCCTGGGTGTGTGTCACCACTGTCAATGAAGCGGTGTGAGTCAATCC 708
Db 618 GGGTTCCTTATCTGCTGTGAAATCTTGCAACTGAAGCCCGAATGAAGTGAATCA 677
Qy 709 GGAACCTTATCTGTCTCCAGAGCTGTGAAGAGATGAACCTTGTGTACTCAGTGTG 768
Db 678 GAGACAGTGTGATCTGCAATCATGATGACATGAATGAATGATACCGCTTGTGTCA 737
Qy 769 GACACCTTATGCTTACAGCTGATTAATCCCACTGTGTATACAGAGTGTG 822
Db 738 GCGTCTTATTCGTTATGACTGGGTGGATTCCGCTGTTAACCCAGGTAG 791

RESULT 14
US-09-880-107-2174/c
Sequence 2174, Application US/09880107
Patent No. US20020142981A1
GENERAL INFORMATION:
APPLICANT: Horne, Darci T.
APPLICANT: Vockley, Joseph G.
APPLICANT: Scherf, Iwe
TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
FILE REFERENCE: 44921-5028-WO
CURRENT APPLICATION NUMBER: US/09/880,107
CURRENT FILING DATE: 2001-06-14
PRIOR APPLICATION NUMBER: US 60/211,379
PRIOR FILING DATE: 2000-06-14
PRIOR APPLICATION NUMBER: US 60/237,054
PRIOR FILING DATE: 2000-10-02
NUMBER OF SEQ ID NOS: 3950
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 2174
LENGTH: 1198
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:

OTHER INFORMATION: Genbank Accession No. US20020142981A1 L20941
US-09-880-107-2174

Query Match 7.9%; Score 175; DB 9; Length 1198;
Best Local Similarity 100.0%; Pred. No. 4.8e-39;
Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2027 CAGCCTTAAATGCTTTTATTCATATAAACTGTGAAGCTAGATGAACCAATTGGAAACA 2086
Db 1198 CAGCCTTAAATGCTTTTATTCATATAAACTGTGAAGCTAGATGAACCAATTGGAAACA 1139
Qy 2087 TTTAATCAGACTGTGATTCAGAGTGGGAACCTTATGTTCTATCTGATCCAGACAG 2146
Db 1138 TTTAATCAGACTGTGATTCAGAGTGGGAACCTTATGTTCTATCTGATCCAGACAG 1079
Qy 2147 CCACACCTTATGATACGCCCCAACTAATGATTTAATTAATACAAATATCTCGTT 2201
Db 1078 CCACACCTTATGATACGCCCCAACTAATGATTTAATTAATACAAATATCTCGTT 1024

RESULT 15
US-10-027-632-134530/c
Sequence 134530, Application US/10027632
Publication No. US20020198371A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 134530
LENGTH: 578
TYPE: DNA
ORGANISM: Human
US-10-027-632-134530

Query Match 7.2%; Score 161.4; DB 13; Length 578;
Best Local Similarity 99.4%; Pred. No. 2.6e-35;
Matches 162; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 582 GCAGGCTTATGACTCCGGCAGAACACAGAGTTGGAGAACTGAGCTTACACACAC 641
Db 183 GCAGGCTTATGACTCCGGCAGAACACAGAGTTGGAGAACTGAGCTTACACACAC 124
Qy 642 ATGTTCTGGGTGCGCTGCTGTGTGTTGGCCAACTGTCAAGAGGCTGTGGAGGT 701
Db 123 ATGTTCTGGGTGCGCTGCTGTGTGTTGGCCAACTGTCAAGAGGCTGTGGAGGT 64
Qy 702 CGAATCCGGAGCCTATCTGCTCCAGAGCTGTGTAACGAGA 744
Db 63 CGAATCCGGAGCCTATCTGCTCCAGAGCTGTGTAACGAGA 21

Search completed: March 26, 2005, 22:07:33
Job time : 1249.65 secs

Db	301	CTACATCCAGCTCATCCCAATTTCTTCTGTCTGAGCTTCACTGACGCTGGTCGAC	360
QY	377	CCGCTGTGGAAACAAGTACGAACCTGCGGTGACCCGACCGCTCACTGAGCTGTGTC	436
Db	361	CCGGCTGTGGAAACAAGTACGAACCTGCGGTGACCCGACCGCTCACTGAGCTGTGTC	420
QY	437	GGGCTTGGTGCAGAGGCAGAGCGAGAGAGGCGGCTCTGCGGCGCACGCTCATCCGCTA	496
Db	421	GGGCTTGGTGCAGAGGCAGAGCGAGAGAGGCGGCTCTGCGGCGCACGCTCATCCGCTA	480
QY	497	CGCCAACTGTGGCAACGTGCTCATCTGCGAGGCTCAGACACCGAGTCTACAGGCGTT	556
Db	481	CGCCAACTGTGGCAACGTGCTCATCTGCGAGGCTCAGACACCGAGTCTACAGGCGTT	540
QY	557	CCCCAGGCCCCAGACCTGTGTGACAGAGGCTTTATGACTCCGGCGAACAACAAGCATTT	616
Db	541	CCCCAGGCCCCAGACCTGTGTGACAGAGGCTTTATGACTCCGGCGAACAACAAGCATTT	600
QY	617	GGAGAAACTGAGCCTTACACAAACATGTTCTGAGGTGACCTGGGTGTGTTGGCCAACT	676
Db	601	GGAGAAACTGAGCCTTACACAAACATGTTCTGAGGTGACCTGGGTGTGTTGGCCAACT	660
QY	677	GTCATTAAGACGCTGTGCTTGAAGGTGAATCCGGGACCTTATCTGCTCCAGAGCTGTCT	736
Db	661	GTCATTAAGACGCTGTGCTTGAAGGTGAATCCGGGACCTTATCTGCTCCAGAGCTGTCT	720
QY	737	GAAAGATGAACACCTTGCGTACTCAGTGTGACACCTGTATGCTTACGACTGGATTAG	796
Db	721	GAAAGATGAACACCTTGCGTACTCAGTGTGACACCTGTATGCTTACGACTGGATTAG	780
QY	797	TATCCACTGTGTGTATACACAGGTGTGACTGTGGCGGTGTACAGCTTCTTCTGACTTGG	856
Db	781	TATCCACTGTGTGTATACACAGGTGTGACTGTGGCGGTGTACAGCTTCTTCTGACTTGG	840
QY	857	TCTAGTTGGGCGGCACTTCTGAAACCCAGCCAAAGCTTACCTGTGCGCTAGACTGACCT	916
Db	841	TCTAGTTGGGCGGCACTTCTGAAACCCAGCCAAAGCTTACCTGTGCGCTAGACTGACCT	900
QY	917	CGTTGTGCCCGTCTTCAAGCTTCTGTGACATGTTCTTATGTTGGCTGGCTGAAGTGGC	976
Db	901	CGTTGTGCCCGTCTTCAAGCTTCTGTGACATGTTCTTATGTTGGCTGGCTGAAGTGGC	960
QY	977	AGAGCAGCTCATCAACCCCTTTGGAGAGATGATGATGATTTTGGACCAACTGGATTGT	1036
Db	961	AGAGCAGCTCATCAACCCCTTTGGAGAGATGATGATGATGATTTTGGACCAACTGGATTGT	1020
QY	1037	CGACAGAAATTTGCAGGTGTCCCTGTTGGCTGTGATGATGATGACACGAGCCTGCGCTCG	1096
Db	1021	CGACAGAAATTTGCAGGTGTCCCTGTTGGCTGTGATGATGATGATGACACGAGCCTGCGCTCG	1080
QY	1097	GATGAGAGCCGGAACATGTACTGGAATTAAGCCCGAGCCACAGCCCCCTTACACAGCTGCTTC	1156
Db	1081	GATGAGAGCCGGAACATGTACTGGAATTAAGCCCGAGCCACAGCCCCCTTACACAGCTGCTTC	1140
QY	1157	CGCCAGATTCGATCGAGCCTCCCTTATGTGGCTCCACCTTCAACATCGAGCCGGAACAAGA	1216
Db	1141	CGCCAGATTCGATCGAGCCTCCCTTATGTGGCTCCACCTTCAACATCGAGCCGGAACAAGA	1200
QY	1217	GGAGATGAGATTCAGGCCCAATCGAGAGACGAGAGAGATGCTTACGCTGTGATCATTTGG	1276
Db	1201	GGAGATGAGATTCAGGCCCAATCGAGAGACGAGAGAGATGCTTACGCTGTGATCATTTGG	1260
QY	1277	CCGCTTCTTAAGGCTGCAAGTCCCAATGATCAACATCTTCCAGAGGCAAACTCAAGACCA	1336
Db	1261	CCGCTTCTTAAGGCTGCAAGTCCCAATGATCAACATCTTCCAGAGGCAAACTCAAGACCA	1320
QY	1337	ACTACTGTGGCCCAAGAGGAATCCCTTCTTCCAGAGAGGCTGTGCCAAAACCAACAAGG	1396
Db	1321	ACTACTGTGGCCCAAGAGGAATCCCTTCTTCCAGAGAGGCTGTGCCAAAACCAACAAGG	1380
QY	1397	AGCCAAACAGAACTTTAGGGGCGCAAGAAACACAAAGCCTTGAAGCTTTAAGCTGTGGA	1456

Db	1381	AGCCAAACAGAAAGCTTATGGGGCCAGAGAAAGACAAACAAAGGCGCTGAAAGCTTAAGCTGTGGA	1440
Qy	1457	CGCCTTCAAGTCTGGGCCACTGTATCAGAGGCCAGGCTACTACAGTGCCTCCCAACAGACGCC	1516
Db	1441	CGCCTTCAAGTCTGGGCCACTGTATCAGAGGCCAGGCTACTACAGTGCCTCCCAACAGACGCC	1500
Qy	1517	CCACAGCCCCACTCCCAAGTGTCTTCCCTTGAACCATCAGCGCGCTCAAGCTTCACAG	1576
Db	1501	CCACAGCCCCACTCCCAAGTGTCTTCCCTTGAACCATCAGCGCGCTCAAGCTTCACAG	1560
Qy	1577	TGTCAACAGCATAGACACCAAGAGCAAAAGCTTAAACATGTGATCTTGGGGCCAAAGAA	1636
Db	1561	TGTCAACAGCATAGACACCAAGAGCAAAAGCTTAAACATGTGATCTTGGGGCCAAAGAA	1620
Qy	1637	AAGTTTGAATGCTCTCTCAGAGACCGATGGGCCCTTGATGAGCAACCCAGAAATATCTCA	1696
Db	1621	AAGTTTGAATGCTCTCTCAGAGACCGATGGGCCCTTGATGAGCAACCCAGAAATATCTCA	1680
Qy	1697	AGTGAAGAGAGAAACTGTGAGTTTAACTGACGAGATTAAGCCAGAAATATCCCGAAATATCA	1756
Db	1681	AGTGAAGAGAGAAACTGTGAGTTTAACTGACGAGATTAAGCCAGAAATATCCCGAAATATCA	1740
Qy	1757	CCTCAAGAACCTTTGGAAACATACCAACCAACATACACATACACTACACTCAAAAGATCAAT	1816
Db	1741	CCTCAAGAACCTTTGGAAACATACCAACCAACATACACATACACTACACTCAAAAGATCAAT	1800
Qy	1817	GGATCCTTATTTGGGCCCTTGGAAAAACAGGATGAGACACATTCCTTAACTGCTTCTTAATG	1876
Db	1801	GGATCCTTATTTGGGCCCTTGGAAAAACAGGATGAGACACATTCCTTAACTGCTTCTTAATG	1860
Qy	1877	GGGATGCTTCCGACAGCAGGTCTCTCACTGTGTGTCACACAGAGAGACATGATCAAGTC	1936
Db	1861	GGGATGCTTCCGACAGCAGGTCTCTCACTGTGTGTCACACAGAGAGACATGATCAAGTC	1920
Qy	1937	ACAGCCATACAGCTGTCTCAACATGAAACATGTCTCTACACAGCCTGAATCAATGATG	1996
Db	1921	ACAGCCATACAGCTGTCTCAACATGAAACATGTCTCTACACAGCCTGAATCAATGATG	1980
Qy	1997	AGCTTAATAGTAAATAATCCCAACATCACTCAAGCCTTAAATGCCCTTAAATCATTAATAAC	2056
Db	1981	AGCTTAATAGTAAATAATCCCAACATCACTCAAGCCTTAAATGCCCTTAAATCATTAATAAC	2040
Qy	2057	TGTGAAAGCTAGACGTGAACCATTGAACATTTAACTCAGACTGTGATTCAGAGTGGGG	2116
Db	2041	TGTGAAAGCTAGACGTGAACCATTGAACATTTAACTCAGACTGTGATTCAGAGTGGGG	2100
Qy	2117	AACCTTAAGTTCATATCGAATCCAAACAGACCAACCTTAAGTATCTGCCAAACTAATG	2176
Db	2101	AACCTTAAGTTCATATCGAATCCAAACAGACCAACCTTAAGTATCTGCCAAACTAATG	2160
Qy	2177	AGTTTAATAAA 2187	
Db	2161	AGTTTAATAAA 2171	

```

; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 16644
; LENGTH: 7609
; TYPE: DNA
; ORGANISM: Human
; OS-09-949-016-16644

```

Query Match	28.8%	Score 642.2;	DB 4;	Length 7609;
Best Local Similarity	99.5%	Pred. No. 4.1e-163;		
Matches 644;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;

Qy	1202	CAGCCTGAACAAAGAGGAGATGGAATTCAGCCCAATACAGAGAGACGAGAGGAATGCTCA	126
Db	7382	CAGCCTGAACAAAGAGGAGATGGAATTCAGCCCAATACAGAGAGAGAGAGGAATGCTCA	7322
Qy	1362	CGCTGGCATCATTTGGCCGCTTCTTAAGGCGCTGCAGTCCCATGATCAACAATCTCTCCAGAGGC	1322
Db	7322	CGCTGGCATCATTTGGCCGCTTCTTAAGGCGCTGCAGTCCCATGATCAACAATCTCTCCAGAGGC	726
Qy	1322	AAACTCAAGGACCAAACTACTGTGGCCCAAGAGGGAATTCCTTCTTCACAGAGGCGCTGCC	138
Db	7262	AAACTCAAGGACCAAACTACTGTGGCCCAAGAGGGAATTCCTTCTTCACAGAGGCGCTGCC	7203
Qy	1382	CAAAAACACAAAGGACGAGCAACAGAAAGTTAGGGGCGCAGGAAGACAACAAGGCGCTGGAA	144
Db	7202	CAAAAACACAAAGGACGAGCAACAGAAAGTTAGGGGCGCAGGAAGACAACAAGGCGCTGGAA	7143
Qy	1442	GCTTAAGGCTGTGAGCGGCTTCAAGTCTGGGCCACTGTATACAGAGGCCAGCTACTACAG	1501
Db	7142	GCTTAAGGCTGTGAGCGGCTTCAAGTCTGGGCCACTGTATACAGAGGCCAGCTACTACTACAG	7083
Qy	1502	TGCCCCAAGAGCGCCCTCAGGCCCACTCCCATGTTCTTCCCGCTTAGAACCATACAGCGCC	156
Db	7082	TGCCCCAAGAGCGCCCTCAGGCCCACTCCCATGTTCTTCCCGCTTAGAACCATACAGCGCC	7023
Qy	1562	GTCAAGCTTCAACAGTGTCAAGGAGTATAGACACCAAGACAAGAAAGCTTAAAGACTGTGAG	1621
Db	7022	GTCAAGCTTCAACAGTGTCAAGGAGTATAGACACCAAGACAAGAAAGCTTAAAGACTGTGAG	6966
Qy	1622	TTCTGGGGGCCAAGAAAGTTTGAATTGCTCTCAGAGAGCGATGGGGCTTGATGAGACA	1681
Db	6962	TTCTGGGGGCCAAGAAAGTTTGAATTGCTCTCAGAGAGCGATGGGGCTTGATGAGACA	6903
Qy	1682	CCCAAGAAATATCTCAAGTGAAGAGGAGAAACGTGGAAGTTTAACTCGACGAGATATGCCAGA	1741
Db	6902	CCCAAGAAATATCTCAAGTGAAGAGGAGAAACGTGGAAGTTTAACTCGACGAGATATGCCAGA	6843
Qy	1742	GATCCCCGAAATCACTCAAGAAACCTTTGGAAACATCACCACCAACATACACACTTAC	1801
Db	6842	GATCCCCGAAATCACTCAAGAAACCTTTGGAAACATCACCACCAACATACACACTTAC	6783
Qy	1802	ACTCAAAAGATCACTGATGATCCTTATATGGGCGTTTGGAAAAACAGGATG 1848	
Db	6782	ACTCAAAAGATCACTGATGATCCTTATATGGGCGTTTGGAAAAACAGGATG 6736	

```

RESULT 3
US-09-949-016-16643
: Sequence 16643, Application US/09949016
: Patent No. 6812339
: GENERAL INFORMATION:
: APPLICANT: VENTER, J. Craig et al.
: TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
: TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREFOR
: FILE REFERENCE: C1001307
: CURRENT APPLICATION NUMBER: US/09/949,016
: CURRENT FILING DATE: 2000-04-14
: PRIOR APPLICATION NUMBER: 60/241,755
: PRIOR FILING DATE: 2000-10-20
: PRIOR APPLICATION NUMBER: 60/237,768
: PRIOR FILING DATE: 2000-10-03
: PRIOR APPLICATION NUMBER: 60/231,498
: PRIOR FILING DATE: 2000-09-08
:

```

```

? NUMBER OF SEQ ID NOS: 207012
? SOFTWARE: PatsSeq for Windows Version 4.0
? SEQ ID NO: 16643
? LENGTH: 18075
? TYPE: DNA
? ORGANISM: Human
US-09-949-016-16643

```

Query Match	28.8%;	Score 642.2;	DB 4;	Length 18075;
Best Local Similarity	99.5%;	Pred. No. 7.1e-163;		
Matches 644;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0

OY	1202	CAGCCTGAAACAAAGAGAGATGAGATGTCACGCCCAATCAGAGAGAGAGAGATGCTCA	1261
Db	13879	CAGCCTGAAACAAAGAGAGATGAGATGTCACGCCCAATCAGAGAGAGAGAGATGCTCA	13938
OY	1262	CGCTGGCATCATTTGGCCGCTTCTCTAGGCTCGACAGTCCCATGATCACAATCTCCAGGGC	1321
Db	13939	CGCTGGCATCATTTGGCCGCTTCTAGCCTCGACAGTCCCATGATCACAATCTCTCCAGGGC	13998
OY	1322	AAACTCAGAGACCAAACTACTGTGGCCCAAGAGGGAATCCCTTCTCCACGAGG3CCTGGC	1381
Db	13999	AAACTCAAGAGACCAAACTACTGTGGCCCAAGAGGGAATCCCTTCTCCACGAGG6CCTGGC	14058
OY	1382	CAAAAACCAAGAGAGACCAACAGAACGTTAGGGGCCAGAGAACAAACAGG6CTGGAA	1441
Db	14059	CAAAAACCAAGAGAGACCAACAGAACGTTAGGGGCCAGAGAACAAACAGG6CTGGAA	14118
OY	1442	GCTTAAGGCTGTGAGCGCCTTCAAGCTGTGGCCCACTGTATCAGAGGCAAGCTTACTACAG	1501
Db	14119	GCTTAAGGCTGTGAGCGCCTTCAAGCTGTGGCCCACTGTATCAGAGGCAAGCTTACTACAG	14178
OY	1502	TGCCCCACAGAGCGCCCTTCA6CCCCACTGCCATGTTCTTCCCTTGAACCATCAGG6CC	1561
Db	14179	TGCCCCACAGAGCGCCCTTCA6CCCCACTGCCATGTTCTTCCCTTGAACCATCAGG6CC	14238
OY	1562	GTCAAAAGCTTCACAGTGTACAGGCACTAGACACAAAGACAAAGGCTTAAAGACTGTGAG	1621
Db	14239	GTCAAAAGCTTCACAGTGTACAGGCACTAGACACAAAGACAAAGGCTTAAAGACTGTGAG	14298
OY	1622	TTCTGGGGCCAAAGAAAGTTTGAATTGCTCTCAGAGCGATGGG6CCTGTATGAGCA	1681
Db	14289	TTCTGGGGCCAAAGAAAGTTTGAATTGCTCTCAGAGCGATGGG6CCTGTATGAGCA	14358
OY	1682	CCCAAGATATCTCAAGTGAGAGAGAAACTGTGAGTTTAACCTGACGATATGCCAGA	1741
Db	14359	CCCAAGATATCTCAAGTGAGAGAGAAACTGTGAGTTTAACCTGACGATATGCCAGA	14418
OY	1742	GATCCCCCAAAATCAGCTCAAGAAACCTTGTGAAACATACACAAACCACTACACTAC	1801
Db	14419	GATCCCCCAAAATCAGCTCAAGAAACCTTGTGAAACATACACAAACCACTACACTAC	14478
OY	1802	ACTCAAGATCATGAGATCTTATTTGGGCTTGTGAAAACAGGATG	1848
Db	14479	ACTCAAGATCATGAGATCTTATTTGGGCTTGTGAAAACAGGATG	14525

```

RESULT 4
US-09-949-016-4902/c
; Sequence 4902, Application US/09949016
; Patent No. 6812339
;
; GENERAL INFORMATION:
;
; APPLICANT: VENTER, J. Craig et al.
;
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
;
; CURRENT APPLICATION NUMBER: US/09/949,016
;
; CURRENT FILING DATE: 2000-04-14
;
; PRIOR APPLICATION NUMBER: 60/241,755
;
; PRIOR FILING DATE: 2000-10-20
;
; PRIOR APPLICATION NUMBER: 60/237,766
;
; PRIOR FILING DATE: 2000-10-03
;
; PRIOR APPLICATION NUMBER: 60/231,498
;

```

;; PRIOR FILING DATE: 2000-09-08
;; NUMBER OF SEQ ID NOS: 207012
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 4902
;; LENGTH: 1462
;; TYPE: DNA
;; ORGANISM: Human
US-09-949-016-4902

Query Match 16.1%; Score 358.4; DB 4; Length 1462;
Best Local Similarity 99.7%; Pred. No. 1.4e-86;
Matches 359; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1842 AGGATGAGACCACTTCTTAACCTGCTTCTTAATGGGATGCTTGGCCAGCCAGGCTCTC 1901
DB 1385 AGGATGAGACCACTTCTTAACCTGCTTCTTAATGGGATGCTTGGCCAGCCAGGCTCTC 1326
QY 1902 ACCTGTGTACACGAGACGAGACCTGATCCAGTCACGCCATACAGCTGTCCACTGA 1961
DB 1325 ACCTGTGTACACGAGACGAGACCTGATCCAGTCACGCCATACAGCTGTCCACTGA 1266
QY 1962 AGAAGTGTCTTACACGAGCTGATCAATGTTAGTTATGATTAATAATCCAGAC 2021
DB 1265 AGAAGTGTCTTACACGAGCTGATCAATGTTAGTTATGATTAATAATCCAGAC 1206
QY 2022 TACTTCAGCCTTTATGCTTCTTTATTCATAAAACCTGGAAGCTAGACTGAACATTGG 2081
DB 1205 TACTTCAGCCTTTATGCTTCTTTATTCATAAAACCTGGAAGCTAGACTGAACATTGG 1146
QY 2082 AAACATTTAACTCACTCTGATTCAGATCGAGTGGGAAACCTTATGTTCTGATCAATCAA 2141
DB 1145 AAACATTTAACTCACTCTGATTCAGATCGAGTGGGAAACCTTATGTTCTGATCAATCAA 1086
QY 2142 GACAGCCACACCTTAGTATCTGCGCCAAACCTAATGATTTATTAATACAAATCTGTT 2201
DB 1085 GACAGCCACACCTTAGTATCTGCGCCAAACCTAATGATTTATTAATACAAATCTGTT 1026

RESULT 5

US-09-949-016-173056
;; Sequence 173056, Application US/09949016
;; Patent No. 6812339
;; GENERAL INFORMATION:
;; APPLICANT: VENTER, J. Craig et al.
;; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;; FILE REFERENCE: CL001307
;; CURRENT APPLICATION NUMBER: US/09/949,016
;; PRIOR FILING DATE: 2000-04-14
;; PRIOR APPLICATION NUMBER: 60/241,755
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/237,768
;; PRIOR FILING DATE: 2000-10-03
;; PRIOR APPLICATION NUMBER: 60/231,498
;; PRIOR FILING DATE: 2000-09-08
;; NUMBER OF SEQ ID NOS: 207012
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 173056
;; LENGTH: 601
;; TYPE: DNA
;; ORGANISM: Human
US-09-949-016-173056

Query Match 8.5%; Score 188.4; DB 4; Length 601;
Best Local Similarity 99.5%; Pred. No. 1e-40;
Matches 189; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1842 AGGATGAGACCACTTCTTAACCTGCTTCTTAATGGGATGCTTGGCCAGCCAGGCTCTC 1901
DB 412 AGGATGAGACCACTTCTTAACCTGCTTCTTAATGGGATGCTTGGCCAGCCAGGCTCTC 471
QY 1902 ACCTGTGTACACGAGACGAGACCTGATCCAGTCACGCCATACAGCTGTCCACTGA 1961

DB 472 ACCTGTGTACACGAGACGAGACCTGATCCAGTCACGCCATACAGCTGTCCACTGA 531
QY 1962 AGAAGTGTCTTACACAGCCTGAATGTTAGTTATGATTAATAATCCAGAC 2021
DB 532 AGAAGTGTCTTACACAGCCTGAATGTTAGTTATGATTAATAATCCAGAC 591
QY 2022 TACTTCAGCC 2031
DB 592 TACTTCAGCC 601

RESULT 6

US-09-949-016-173051
;; Sequence 173051, Application US/09949016
;; Patent No. 6812339
;; GENERAL INFORMATION:
;; APPLICANT: VENTER, J. Craig et al.
;; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;; FILE REFERENCE: CL001307
;; CURRENT APPLICATION NUMBER: US/09/949,016
;; PRIOR FILING DATE: 2000-04-14
;; PRIOR APPLICATION NUMBER: 60/241,755
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/237,768
;; PRIOR FILING DATE: 2000-10-03
;; PRIOR APPLICATION NUMBER: 60/231,498
;; PRIOR FILING DATE: 2000-09-08
;; NUMBER OF SEQ ID NOS: 207012
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 173051
;; LENGTH: 601
;; TYPE: DNA
;; ORGANISM: Human
US-09-949-016-173051

Query Match 7.2%; Score 161.4; DB 4; Length 601;
Best Local Similarity 99.4%; Pred. No. 2.1e-33;
Matches 162; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 582 GCAGGCTTTATGACTCCGGGAGACACAGCACTTGGAGAAATGAGCTTACACACAC 641
DB 321 GCAGGCTTTATGACTCCGGGAGACACAGCACTTGGAGAAATGAGCTTACACACAC 380
QY 642 ATGTTCCGGGTGCGCTGGGGTGTGTTGCCAACCTGTCATGAAGGGTGGTGGAGGT 701
DB 381 ATGTTCCGGGTGCGCTGGGGTGTGTTGCCAACCTGTCATGAAGGGTGGTGGAGGT 440
QY 702 CGAATCCGGGACCTTATCTGCTCCAGAGCTGCTGAACGAGA 744
DB 441 CGAATCCGGGACCTTATCTGCTCCAGAGCTGCTGAACGAGA 483

RESULT 7

US-09-949-016-173052
;; Sequence 173052, Application US/09949016
;; Patent No. 6812339
;; GENERAL INFORMATION:
;; APPLICANT: VENTER, J. Craig et al.
;; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;; FILE REFERENCE: CL001307
;; CURRENT APPLICATION NUMBER: US/09/949,016
;; PRIOR FILING DATE: 2000-04-14
;; PRIOR APPLICATION NUMBER: 60/241,755
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/237,768
;; PRIOR FILING DATE: 2000-10-03
;; PRIOR APPLICATION NUMBER: 60/231,498
;; PRIOR FILING DATE: 2000-09-08
;; NUMBER OF SEQ ID NOS: 207012
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 173052

LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-173052

Query Match 7.2%; Score 161.4; DB 4; Length 601;
Best Local Similarity 99.4%; Pred. No. 2.1e-33;
Matches 162; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 582 GCAGGCTTTATGACTCCGGCAGAACACAGAGATTGGAGAACTGAGCTTACACACAC 641
DB 30 GCAGGCTTTATGACTCCGGCAGAACACAGAGATTGGAGAACTGAGCTTACACACAC 89
QY 642 ATGTTCTGGGTCCTGGGTCCTGGGTCCTGGGTCCTGGGTCCTGGGTCCTGGGTCCT 701
DB 90 ATGTTCTGGGTCCTGGGTCCTGGGTCCTGGGTCCTGGGTCCTGGGTCCTGGGTCCT 149
QY 702 CGAATCCGGGACCTATCTCTCTCCAGAGCTGCTGAACGAGA 744
DB 150 CGAATCCGGGACCTATCTCTCTCCAGAGCTGCTGAACGAGA 192

RESULT 8
US-09-949-016-173045
Sequence 173045, Application US/0949016
Patent No. 6812339

GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: C1001307
CURRENT APPLICATION NUMBER: US/09/949.016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 173045
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-173045

Query Match 5.7%; Score 128; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 2.5e-24;
Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 69 GACCAAGCCCACTGCTGACGCCCACTGCTGCGCATGACCATCACTTACAAAGCCAA 128
DB 474 GACCAAGCCCACTGCTGACGCCCACTGCTGCGCATGACCATCACTTACAAAGCCAA 533
QY 129 GTGGCTATGCGCGCTTGGCTCTTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 188
DB 534 GTGGCTATGCGCGCTTGGCTCTTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 593
QY 189 TACAAGCT 196
DB 594 TACAAGCT 601

RESULT 9
US-09-270-767-735
Sequence 735, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767

CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 735
LENGTH: 419
TYPE: DNA
ORGANISM: Drosophila melanogaster
US-09-270-767-735

Query Match 4.6%; Score 102.4; DB 4; Length 419;
Best Local Similarity 62.9%; Pred. No. 1.7e-17;
Matches 176; Conservative 0; Mismatches 101; Indels 3; Gaps 1;

QY 944 GTTCTTCTTCAATGTTGGCTGGCTGGAAGTGGGAGAGAGCTCATCAACCCCTTTGAGA 1003
DB 1 GTTCTTCTTCAATGTTGGCTGGCTGGAAGTGGGAGAGAGCTCATCAACCCCTTTGAGA 60
QY 1004 GGATGATGATGATTTTGAAGCAACCTGATTTGACAGAAATTTGACAGGTGCTGTT 1063
DB 61 AGACGATGATGATTTTGAAGTCACTGATGATGATGATGATGATGATGATGATGATGAT 120
QY 1064 GCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1123
DB 121 GATGTCGACGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
QY 1124 GCCCGAGCC--ACAGCCCCCTACACAGCTGCTTCCGCCAGTTCGTCGAGCTCCTT 1180
DB 181 GGTGTTCCCAACGAGCTGCTTCAACATGATGCTGCGAATGATTTCCGGAGAAATCATCC 240
QY 1181 TATGGGCTCCACTTCAACATGATGATGATGATGATGATGATGATGATGATGATGAT 1220
DB 241 AGAGCCGTCACCTGCAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 280

RESULT 10
US-09-270-767-16017
Sequence 16017, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 16017
LENGTH: 419
TYPE: DNA
ORGANISM: Drosophila melanogaster
US-09-270-767-16017

Query Match 4.6%; Score 102.4; DB 4; Length 419;
Best Local Similarity 62.9%; Pred. No. 1.7e-17;
Matches 176; Conservative 0; Mismatches 101; Indels 3; Gaps 1;

QY 944 GTTCTTCTTCAATGTTGGCTGGCTGGAAGTGGGAGAGAGCTCATCAACCCCTTTGAGA 1003
DB 1 GTTCTTCTTCAATGTTGGCTGGCTGGAAGTGGGAGAGAGCTCATCAACCCCTTTGAGA 60
QY 1004 GGATGATGATGATTTTGAAGCAACCTGATTTGACAGAAATTTGACAGGTGCTGTT 1063
DB 61 AGACGATGATGATTTTGAAGTCACTGATGATGATGATGATGATGATGATGATGATGAT 120
QY 1064 GCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1123
DB 121 GATGTCGACGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
QY 1124 GCCCGAGCC--ACAGCCCCCTTACACAGCTGCTTCCGCCAGTTCGTCGAGCTCCTT 1180
DB 181 GGTGTTCCCAACGAGCTGCTTCAACATGATGATGATGATGATGATGATGATGATGATGAT 240
QY 1181 TATGGGCTCCACTTCAACATGATGATGATGATGATGATGATGATGATGATGATGAT 1220

Db 241 AGAGCGTCCACTGCCAAGATGAGGTGCCCAAGATGCG 280

```

RESULT 11
; US-09-949-016-173053
; Sequence 173053, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 173053
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-173053

```

Query Match	3.2%	Score 71.4;	DB 4;	Length 601;
Best Local Similarity	92.6%	Pred. No. 5.5e-09;		
Matches 75; Conservative	0;	Mismatches 6;	Indels 0;	Gaps 0

QY	756	CGATCCAGGTCGGAGAACCTGATATGCTTACGACTGATATTAGTCCCATCTGATGTATACA	815
Dd	1	CGATCCAGGTCGGAGAACCTGATATGCTTACGACTGATATTAGTCCCATCTGATGTATACA	60
QY	816	CAGCTGCTGACTGTGCGGTG	836
Dd	61	CAGGTGAGACTAGGCTGGTG	81

```

; RESULT 12
; US-09-949-016-173054
; Sequence 173054, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 173054
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-173054

```

Query Match	3.2%	Score 70.4;	DB 4;	Length 601;
Best Local Similarity	92.5%	Pred. No. 1e-08;		
Matches 74;	Conservative 0;	Mismatches 6;	Indels 0;	Gaps 0;

757 G T A C T C A G T G T G A C A C C T G T A T G C C T A C G A C T G A T T A G T A C C C A C T G T G T A T A C A C 816

Db 1 GTACTAGTGTGGACACTGTATAGCTACGACTGATTTAGTATCCACTGGTGTATACAC 60
Oy 817 AGGCTGTACTGTGGCGGTG 836
Db 61 AGGTAGAGACTAGGCTGGTT 80

```

RESULT 13
US-09-949-016-173053/c
; Sequence 173053, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C0001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 173053
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-173053

```

Query Match 3.0%; Score 66; DB 4; Length 601;
Best Local Similarity 79.6%; Pred. No. 1.6e-07;
Matches 78; Conservative 0; Mismatches 20; Indels 0; Gaps 0

QY 186 CAGCTTCCATATGGGAGATGCTTGCCAGCCAGAGTCTCACTGTGTGTAACCCAGCAGAG 19
Db 103 CAGTTTCCCAAAAAGGCGAGCCTCACCAGCTATGCTCTTCACTGTGTATACACCAGTGGG 44
QY 1924 CACTGATCCAGTACAGCCCATATACAGCTGTCCACACTGA 1961
Db 43 TACTATTCAGTCCGTAGGACATACAGGTGTCCCACTGA 6

```

US-09-949-016-173054/c
/ Sequence 173054, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ WITH INFECTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: C1001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 173054
/ LENGTH: 601
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-173054

```

Query Match	3.0%;	Score 66;	DB 4;	Length 601;
Best Local Similarity	79.6%;	Pred. No. 1.6e-07;		
Matches 78;	Conservative	0;	Mismatches 20;	Indels 0;
			Gaps	0;

Oy 1864 CTGCTTCTTAATGGGATGCTTCGCCAGCAGCTCCTCACTGTGTGTACACACAGAGA 1923
 Db 102 CAGTTTCCCAAAAGGGCAGCCTCACCAGCCTAGTCTCCTCACTGTGTATACACAGTGGGA 43
 Oy 1924 CACTGATCCAGTCACAGCCATACAGCTGTTCACACTGA 1961
 Db 42 TACTAATCCAGTCGTAGGCATACAGTGTCCACACTGA 5

RESULT 15
 US-09-949-016-16643/c
 ; Sequence 16643, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; FILE REFERENCE: CLO01307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO: 16643
 ; LENGTH: 18075
 ; TYPE: DNA
 ; ORGANISM: Human
 US-09-949-016-16643

Query Match 3.0%; Score 66; DB 4; Length 18075;
 Best Local Similarity 79.6%; Pred. No. 1.3e-06;
 Matches 78; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

Oy 1864 CTGCTTCTTAATGGGATGCTTCGCCAGCAGCTCCTCACTGTGTGTACACACAGAGA 1923
 Db 9126 CAGTTTCCCAAAAGGGCAGCCTCACCAGCCTAGTCTCCTCACTGTGTATACACAGTGGGA 9067
 Oy 1924 CACTGATCCAGTCACAGCCATACAGCTGTTCACACTGA 1961
 Db 9066 TACTAATCCAGTCGTAGGCATACAGTGTCCACACTGA 9029

Search completed: March 26, 2005, 13:49:34
 Job time : 357.338 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_p2n model

Run on: March 23, 2005, 06:37:09 ; Search time 7468.71 Seconds
(without alignments)
3795.340 Million cell updates/sec

Title: US-09-622-964A-3

Perfect score: 3120

Sequence: 1 MTTTYSQVNNARLGSFSL.....TLKHMDPYALENRDAHS 585

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=cgnt_1/USPTO.spool/US09622964/rnatc_22032005_062218_26369/app_query.fasta_1.2069
-DB=GenEmbl -QFMT=faaetap -SUFFIX=xge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=bloum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=spct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09622964.@CGN_1.1_1531.@rnatc_22032005_062218_26369 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -Delop=6 -Delext=7

Database :

GenEmbl:*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_pn:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	3116	99.9	2229	6 BD136720
2	3110	99.7	1758	6 AX745964
3	3110	99.7	1758	9 AY515704
4	3110	99.7	2171	9 AF073501

5	3110	99.7	2210	9 AF057169	AF057169 Homo sapi
6	3083	98.8	2170	6 CQ729633	CQ729633 Sequence
7	3047.5	97.7	2429	6 BD136721	BD136721 Best's ma
8	3041.5	97.5	2420	9 AF057170	AF057170 Homo sapi
9	3001	96.2	2187	9 AY357925	AY357925 Macaca fa
10	2817	90.3	2435	9 BC041664	BC041664 Homo sapi
11	2672.5	85.7	2441	9 BC015220	BC015220 Homo sapi
12	1914	61.3	2035	10 BC079048	BC079048 Rattus no
13	1899	60.9	1916	6 BD136744	BD136744 Best's ma
14	1896	60.8	1904	10 AY450427	AY450427 Mus muscu
15	1701	54.5	1289	4 AY064707	AY064707 Sus scrof
16	1655.5	53.1	112309	4 AC003025	AC003025 Human Chr
17	1652.5	53.0	142092	9 AF139813	AF139813 Homo sapi
18	1652.5	53.0	196080	9 AC004228	AC004228 Homo sapi
19	1645.5	52.7	163024	9 AP006260	AP006260 Homo sapi
20	1645.5	52.7	163915	2 AC087451	AC087451 Homo sapi
21	1645.5	52.7	166867	9 AP003733	AP003733 Homo sapi
22	1636.5	52.5	133683	2 AC084857	AC084857 Homo sapi
23	1615	51.8	16125	6 BD136719	BD136719 Best's ma
24	1571	50.4	1263	6 BD132659	BD132659 Secreted
25	1433	45.9	160169	2 AC051664	AC051664 Homo sapi
26	1358.5	43.5	1898	5 BC043854	BC043854 Xenopus l
27	1358.5	43.5	1921	5 AY273825	AY273825 Xenopus l
28	1348	43.2	1882	5 CR760914	CR760914 Xenopus t
29	1348	43.2	2021	5 BC061379	BC061379 Xenopus t
30	1343	43.0	2719	10 AY450426	AY450426 Mus muscu
31	1337.5	42.9	2064	5 BC084229	BC084229 Xenopus l
32	1332	42.7	2007	9 AY515706	AY515706 Homo sapi
33	1332	42.7	2500	6 AX834175	AX834175 Sequence
34	1332	42.7	2500	9 AK096459	AK096459 Homo sapi
35	1329.5	42.6	1650	5 AY273826	AY273826 Xenopus l
36	1311.5	42.0	2028	6 AY745968	AY745968 Sequence
37	1296	41.5	1506	9 AF440758	AF440758 Homo sapi
38	1286.5	41.2	1956	10 BC031186	BC031186 Mus muscu
39	1286.5	41.2	1957	10 BC019528	BC019528 Mus muscu
40	1286.5	41.2	1992	10 BC036157	BC036157 Mus muscu
41	1286.5	41.2	2028	10 BC036163	BC036163 Mus muscu
42	1285.5	41.2	1954	10 AY450428	AY450428 Mus muscu
43	1284	41.2	1908	9 AF440756	AF440756 Homo sapi
44	1282	41.1	2137	6 AX745966	AX745966 Sequence
45	1281.5	41.1	1530	6 CQ731444	CQ731444 Sequence

ALIGNMENTS

RESULT 1	BD136720	2229 bp	DNA	linear	PAT 18-SEP-2002
LOCUS	Best's macular dystrophy gene.				
DEFINITION	BD136720				
ACCESSION	BD136720.1 GI:23231665				
VERSION	JP 2002504559-A/2.				
KEYWORDS	UP 2002504559-A/2.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
TITLE	1 (bases 1 to 2229)				
JOURNAL	Petrunkhin,K., Caskey,T.C., Metzker,M. and Wadelius,C.				
COMMENT	Best's macular dystrophy gene				
	Patent: JP 2002504559-A 2 12-FEB-2002;				
	MERCK & CO INC, CLABS WADELIIUS				
	OS Homo sapiens (human)				
	PN JP 2002504559-A/2				
	PD 12-FEB-2002				
	PR 22-FEB-1999 JP 2000533447				
	PR 25-FEB-1998 US 60/075941,18-DEC-1998 US 60/112926 PI				
	KONSTANTIN PETRUNKHIN,THOMAS C CASKEY,MICHAEL METZKER,CLABS PI				
	WADLIUS				
	PC C07K16/18,C07K14/47,C12N5/10,C12N15/09,C12P19/34,C12Q1/68// PC				
	C12P21/08,				
	PC C12N5/00,C12N15/00				
	CC Best's macular dystrophy gene				
	FH Key Location/Qualifiers				

FT source 1..2229
FT /organism='Homo sapiens (human)'.
FEATURES
source location/Qualifiers
1..2229
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 6 74e-267 Length: 2229
Score: 316.00 Matches: 584
Percent Similarity: 99.83% Conservative: 0
Best Local Similarity: 99.83% Mismatches: 1
Query Match: 99.87% Indels: 0
DB: Gaps: 0
US-09-622-964a-3 (1-585) x BD136720 (1-2229)
QY 1 MetThrIleThrTyThrSerGlnValAlaAsnAlaArgLeuGlySerPheSerArgLeu 20
DB 105 ATGACCATCATTAACAAGCCAGTGGCTAATGCCCGTTAGGCTCTTCCCGCCTG 164
QY 21 LeuLeuCysThrArgGlySerIleTyTyIysLeuLeuTyGlyGluPheLeuIlePheLeu 40
DB 165 CTGCTGTCTGGCGGGGAGACATCAACAAGCTGCTAATGCGAGTTCTTAATCTTCTG 224
QY 41 LeuCysTyTyTlIleIleArPheIleTyArgLeuAlaLeuThGluGluGlnLeu 60
DB 225 CTCTGTACTATCATATCCGCTTTATTATTAGCTGGCCCTCAGGAAGAACACAGCTG 284
QY 61 MetPheGluIysLeuThrLeuTyCysAspSerTyIleGlnLeuIleProIleSerPhe 80
DB 285 ATGTTTGAGAACTGACTCTGTATTGGACAGCTCATCCAGCTCATCCCATTTCCCTC 344
QY 81 ValLeuGlyPheTyValThrIleuValAlaThrArgTPTTPAsnGlnTyGluAsnLeu 100
DB 345 GTCGCTGGCTTTAGCTGAAGCTGGTGGTGGACCCGCTGGTGAACAGTACAGAACTG 404
QY 101 ProTPProAspArgLeuMetSerLeuValSerGlyPheValGluGlyIysAspGluGln 120
DB 405 CGGTGGCCGACCGGCTCATAGAGCTGGTGGCTTCGTGAAAGGACAGAGCAAGCA 464
QY 121 SerArgLeuLeuArgArgThrLeuIleArgTyValAsnLeuGlyAsnValLeuIleLeu 140
DB 465 GAGCGGCTGTCGGCGGACGCTCATTCGCTACGCAACCTGGGCAACGTTGCTCATCCTG 524
QY 141 ArgSerValSerThrAlaValTyTlySarPheProSerAlaGlnIleuValGlnAla 160
DB 525 CGGAGCGTCAGCACCGGCTTCAAGGCTTCCCGAGCGCCAGCACTGGTGCAAGCA 584
QY 161 GlyPheMetThrProAlaGlnIleIysGlnLeuGluIysLeuSerLeuProIleAsnMet 180
DB 585 GCGTTTATGACTCCGGCAGAACACAGAGCTTGGAGAACTGAGCCATCCACACACATG 644
QY 181 PheTrpValProTPValTTPPheAlaAsnLeuSerMetIysValATTPLeuGlyIysArg 200
DB 645 TTCGGGTGCTCTGGGTGGTGGTGGCAACCTGTCAATGAAAGGCTGGTGGAGTCA 704
QY 201 IleArgAspProIleLeuLeuGlnSerIleuAsnGlnuMetAsnThrLeuArgThGln 220
DB 705 ATCCGGGACCTATCCGCTCCAGAGCTGCTGACAGATGAACACCTTGGCGTACTCAG 764
QY 221 CysGlyIleIleuTyValAlaTyAspTrpIleSerIleProLeuValTyTrhGlnValAl 240
DB 765 TGTGGACACCTGTATGCTACAGACTGATTATCCACCTGGGTATACACAGGTGGTG 824
QY 241 ThrValAlaValTySerPhePheLeuThrCysIleuValAlaTyArgGlnPheLeuAsnPro 260
DB 825 ACTGTGGCGGTGACAGCTTCTTCTGACTGTCTGAGTGGGGGCAAGTTTCGAACCCA 884
QY 261 AlaIysAlaTyProGlyIleGluLeuAspLeuValValProValPheThrPheLeuGln 280

DB 885 GCCAAGGCTTACCTGGCGATGAGCTGACCTCGTTGTGCCCGTCTCACGTTCTTCAG 944
QY 281 PhePhePheTyValGlyTyPheIysValAlaGluGlnLeuIleAsnProPheGlyGlu 300
DB 945 TTCCTCTTCTATGTTGGCTGGCTGAGAGTGGCAGAGAGCTCATCAACCTTTGGAGAG 1004
QY 301 AspAspAspAspPheGluThrAsnTrpIleValAlaAspArgAsnLeuGlnValSerLeu 320
DB 1005 GATGATGATGATTTTGAACCAACTGATTGTGCACAGGAATTTGACAGGTGTCCCTGTTG 1064
QY 321 AlaValAspGluMetHisGlnAspLeuProArgMetGluProAspMetTyTrpAsnIys 340
DB 1065 GCTGTGATGATGATGACACAGAGACTCTCCGATGAGCCGACATGTACTGGAAATAG 1124
QY 341 ProGluProGlnProProTyTrhAlaAlaSerAlaGlnPheArgArgAlaSerPheMet 360
DB 1125 CCCAGGACAGCCCCCTTACAGAGCTGCTCCGCCCAAGTTCGTCGAGCTCTTATATG 1184
QY 361 GlySerThrPheAsnIleSerLeuAsnIysGluGluMetGluPheGlnProAsnGlnIu 380
DB 1185 GCTTCCACTTCAACATCACTGACCTGAACAGAGAGATGAGTTCCAGCCCAATCAGAG 1244
QY 381 AspGluGluAspAlaHisAlaGlyIleIleGlyArgPheLeuGlyLeuGlnSerHisAsp 400
DB 1245 GACGAGGAGGATGCTACGCTGGCATCATTTGGCCGCTTCTTAGGCTTCGACGTCCATGAT 1304
QY 401 HisHisProProArgAlaAsnSerArgThrIysLeuLeuTrpProIysArgIleuSerLeu 420
DB 1305 CACCATCTTCCAGGGGCAAACTCAGAGACCAAACTAGTGGCCCAAGAGGAAATCCCTT 1364
QY 421 LeuHisGluGlyLeuProIysAsnHisIlyAlaAlaIysGluAsnValArgIlyGlnIu 440
DB 1365 CTCACAGAGGCTTGCCTCCAAAACCAAGAGCAGCCAAACAGAACTTAGGGGCCAGAA 1424
QY 441 AspAsnIysAlaTrpIysLeuIysAlaValAlaPheIysSerGlyProLeuTyTrgln 460
DB 1425 GACAAAGAGCCCTGAAGCTTAAGCTGTGAGAGCTTCAAGCTTCCAGCTGATCAG 1484
QY 461 ArgProGlyTyTrpSerAlaProGlnThrProLeuSerProThrProMetPhePhePro 480
DB 1485 AGGCGAGGCTACTACAGTCCCAAGAGAGCCCTCAGGCCCACTCCCAATGTTCTTCCC 1544
QY 481 LeuGluProSerAlaProSerIysLeuHisSerValThGlyIleAspThrIysAspIys 500
DB 1545 CTAGAACATACAGCGCGCTCAAAAGCTTCACTGATCAGAGCATGACCAACCAACAA 1604
QY 501 SerLeuIysThrValSerSerGlyAlaIysIysSerPheGluLeuLeuSerGluSerAsp 520
DB 1605 AGCTTAAAGACTGTGAGTTCTGGGGCCAAAGAAAGTTTGAATTGCTCTCAGAGAGCAT 1664
QY 521 GlyAlaLeuMetGluHisProGluValSerGlnValArgArgIysThrValGluPheAsn 540
DB 1665 GGGGCTTATGAGAGCAACCAAGATATCTCAAGTGAAGAGAAACCTGTGAGTTTAC 1724
QY 541 LeuThrAspMetProGluIleProGluAsnHisIleuIysGluProLeuGluGlnSerPro 560
DB 1725 CTGACGATATCCAGAGATCCCGGAAATCACTCAAGAAACCTTTGGAACATACCA 1784
QY 561 ThrAsnIleHisThrThrLeuIysAspHisMetAspProTyTrpAlaLeuGluAsnArg 580
DB 1785 ACCAACAATACACATCACTCAAAAGATCAATGATCTTATTTGGGCTTGGAAAAACAG 1844
QY 581 AspGluAlaHisSer 585
DB 1845 GATGAAGCAATTCC 1859
RESULT 2
AX745964 1758 bp DNA linear PAT 14-MAY-2003
LOCUS AX745964
DEFINITION Sequence 1 from Patent WO03030922.
ACCESSION AX745964
VERSION AX745964.1 GI:30724619
KEYWORDS

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1
AUTHORS Steuernagel, A., Bruenner, G., Fritsch, R., Eulenberg, K. and
Clossak, T.
TITLE Bestrophin and bestrophin homologous proteins involved in the
regulation of energy homeostasis
JOURNAL Patent: WO 03030922-A 1 17-APR-2003;
Deutscher Patentantrag fuer Entwicklungsbioologische Forschung
(DE)
FEATURES
source location/Qualifiers
1..1758
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 1,68e-266 Length: 1758
Score: 3110.00 Matches: 583
Percent Similarity: 99.66% Conservative: 0
Best Local Similarity: 99.66% Mismatches: 2
Query Match: 99.68% Indels: 0
DB: 6 Gaps: 0
US-09-622-964A-3 (1-585) x AX745964 (1-1758)
Qy 1 MetThrIleThrYrThrSerGlnValAlaAsnAlaArgLeuGlySerPheSerArgLeu 20
Db 1 ATGACATCACTTACACAGCAAGTGGCTAATCCCGCTTAGGCTCTTCCCGCTG 60
Qy 21 LeuLeuCystrPArgIySerIleTyrluLeuLeuTyrglyGluPheLeuIlePheLeu 40
Db 61 CTGCTGTGCTGGCCGGGACAGCATCTACAGCTGCTAATATGCGAGTCTTAATCTTCTG 120
Qy 41 LeuCystrYrTrIleIleArgPheIleTyrlArgLeuAlaLeuThrGluGlnGlnLeu 60
Db 121 CTGCTGACTACATCATCCGCTTAATTAATAGGCTGGCCCTCACGAAAGAACACAGCTG 180
Qy 61 MetPheGluLeuLeuThrLeuTyrlCytrAspSerTyrlIleGlnLeuIleProIleSerPhe 80
Db 181 ATGTTTGAAGAACTGACTCTGTATTTGGCAGCTACATCCAGCTCATTCCTTCCTTC 240
Qy 81 ValLeuGlyPheTyrlValThrLeuValValThrArgTrpPheAsnGlnTyrglyAsnLeu 100
Db 241 GTGCTGGGCTTCTACGTGACGCTGCTGCTGACCCGCTGGTGAACAGTACGAAACCTG 300
Qy 101 ProTrpProAspArgLeuMetSerLeuValSerGlyPheValGlyGlyValAspGlyGln 120
Db 301 CCGTGGCCCAACCCCTCATGAGCCTGTGTCTGGGCTTCTCGAAGGCAAGGACGACCAA 360
Qy 121 SerArgLeuLeuArgArgThrLeuIleArgTyrlaAsnLeuGlyAsnValLeuIleLeu 140
Db 361 GGCGGGCTGCTGGCGGCGGACGCTCATCCGCTAACCCGCAACCTGGGCAAGTCTCATC 420
Qy 141 ArgSerValSerThrAlaValTyrluArgPheProSerAlaGlnIleValuValGlnAla 160
Db 421 CGCAGCCTCAGCAGCCGAGCTTCAAGGCTTCCCGAGCGCCGACGACCTGGTCAAGCA 480
Qy 161 GlyPheMetThrProAlaGlnIleValGlyGlnLeuGlyValSerLeuSerProIleAsnMet 180
Db 481 GGGCTTTATGACTCGGCGAAGACACAGAGTGGAGAACTGAACCTTACCAACAACATG 540
Qy 181 PheTrpValProTrpValTrpPheAlaAsnLeuSerMetLeuValATrPLeuGlyIArg 200
Db 541 TTCTCGGCTGCTGGCTGTGTGTGTTGCCAAGCTTCAATGAAGGCTGCTTGGAGTCTGA 600
Qy 201 IleArgAspProIleLeuLeuGlnSerLeuLeuAsnGluMetAsnThrLeuArgThrGln 220
Db 601 ATCCGGAACCTATCTGCTCCAGAGCCTGCTGAACGAGATGAACACCTTCCGTAATCAG 660

Qy 221 CysGlyHisLeuTyrlaValAspTrpIleSerIleProLeuValTyrlThrGlnValVal 240
Db 661 TGTGGACACCTGATGCTTACGACTGGATTTAGATCCACAGTGGTGTATACAGAGTGCTG 720
Qy 241 ThrValAlaValTyrlSerPhePheLeuThrCysLeuValGlyArgGlnPheLeuAsnPro 260
Db 721 ACTGTGGCGGTGATACACTTCTTCTGACTTGTCTAGTGTGGCGGCGAGTTCTTCAACCCA 780
Qy 261 AluValAlaTyrlProGlyHisGlyLeuAspLeuValValProValPheThrPheLeuGln 280
Db 781 GCCAAGGCTTACCTGGCCATGAGCTGACCTCTGTGGCCGCTTCTTCAAGCTTCTGAG 840
Qy 281 PhePhePheTyrlValGlyTrpLeuValAlaGlnGlnLeuIleAsnProPheGlyGlu 300
Db 841 TTCTTCTTCTATGTGGCTGGCTGTAAGTGGCAGAGAGCATCAACCCCTTTGGAGAG 900
Qy 301 AspAspAspAspPheGluThrAsnTrpIleValAspArgAsnLeuGlnValSerLeuLeu 320
Db 901 GATGATGATGATTTTGAACCAACTGGAATTGTGACAGAGAAATTTGACAGTGTCTCCCTG 960
Qy 321 AlaValAspGluMetHisGlnAspLeuProArgMetGluProAspMetTyrlTrpAsnLys 340
Db 961 GCTGTGATGATGATGACACAGAGACTGCTCGGATGGAGCCGACATGTACTGGAATTAAG 1020
Qy 341 ProGluProGluProProTyrlThrAlaAlaSerAlaGlnPheArgArgAlaSerPheMet 360
Db 1021 CCCGAGCCACAGCCCTTACAGAGCTGTCCGCCAAGTTCCGTGAGCTTCTTTATG 1080
Qy 361 GlySerThrPheAsnIleSerLeuAsnLysGlyGluMetGluPheGlnProAsnGlnGlu 380
Db 1081 GGCTCCACCTTCAACATCACTGACTGTAACAAGAGAGATGAGTCCAGCCCAATCAGAG 1140
Qy 381 AspGluGluAspAlaHisAlaGlyIleIleGlyArgPheLeuGlyLeuGlnSerHisAsp 400
Db 1141 GACAGAGAGATGCTCACCGTGCATATGGCGCTTCTTACAGCTGACAGTCCCATGAT 1200
Qy 401 HisIleProProAspAlaAsnSerArgThrIleLeuLeuTyrlProLysArgGluSerLeu 420
Db 1201 CACCATCTTCCAGGCGGAACCTCAAGSACCAAACTACTGTGGCCCAAGAGGAATCCCTT 1260
Qy 421 LeuHisGlyGlyLeuProLysAsnHisLysAlaAlaLysGlnAsnValArgGlyGlnGlu 440
Db 1261 CTCACAGAGGCTGCGCCAAAAACCAAGAGCAGCCAAACAGAGCTTAAAGGGCCAGAGAA 1320
Qy 441 AspAsnLysAlaTrpLysLeuValValAlaAspAlaPheLysSerGlyProLeuTyrlGln 460
Db 1321 GACAAACAAGGCTTGAAGCTTAAAGCTGTGAACGCTTCAAGTCTGCCCACTGTATCTAG 1380
Qy 461 ArgProGlyTyrlTyrlSerAlaProGlnThrProLeuSerProThrProMetPhePhePro 480
Db 1381 AGGCCAGGCTACTACAGTGTGCCCAAGACACCCCTCAGGCCCACTCCCATGTCTTCCCC 1440
Qy 481 LeuGluProSerAlaProSerLysLeuHisSerValThrGlyIleAspThrLysAspLys 500
Db 1441 CTGAACCATCAGAGCGCTCAAAAGCTTCAACAGTGTCAAGGCAATAGACACCAAGACAAA 1500
Qy 501 SerLeuLysThrValSerSerGlyAlaLysLysSerPheGlyLeuLeuSerGluSerAsp 520
Db 1501 AGCTTAAAGACTGAGATTTCTGGGGCCAAAGAAAGTGTGAATTGCTCTCAGAGAGGAT 1560
Qy 521 GlyAlaLeuMetGluHisProGluValSerGlnValArgArgLysThrValGluPheAsn 540
Db 1561 GGGGCTTGTATGAGACCCCAAGAGTATCTCAAGTGAAGAGAAACTGTGGAGTTTAAAC 1620
Qy 541 LeuThrAspMetProGluIleProGluAsnHisLysLeuGluProLeuGluGlnSerPro 560
Db 1621 CTGACGATATGCAAGAGATCCCGAATAATCACTTCAAGAACTTTGAAACATTCACCA 1680
Qy 561 ThrAsnIleHisThrThrLeuLysAspHisLysAspProTyrlTrpAlaLeuGluAsnArg 580
Db 1681 ACCAACTATCACTTACACTCAACAAAGATCAATGATGATCTTAATGGGCTTGGAAAAACAG 1740
Qy 581 AspGluAlaHisSer 585

Db 1741 GATGAGACATTC 1755

|||||

RESULT 3
LOCUS AY515704 1758 bp mRNA linear PRI 28-JAN-2004
DEFINITION Homo sapiens beactrophin 1 mRNA, complete cds.
ACCESSION AY515704
VERSION AY515704.1 GI:41216872
KEYWORDS

SOURCE
Homo sapiens (human)

ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE
AUTHORS Sun,H., Tsunenari,T., Yau,K.W. and Nathans,J.
TITLE The vitelliform macular dystrophy protein defines a new family of chloride channels
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (6), 4008-4013 (2002)
MEDLINE 21902086
PUBMED 11904445

REFERENCE
AUTHORS 2 (bases 1 to 1758)
Tsunenari,T., Sun,H., Williams,J., Cahill,H., Smallwood,P.,
Yau,K.W. and Nathans,J.
TITLE Structure-function analysis of the beactrophin family of anion channels
J. Biol. Chem. 278 (42), 41114-41125 (2003)
MEDLINE 22917481
PUBMED 12907679

REFERENCE
AUTHORS 3 (bases 1 to 1758)
Sun,H., Tsunenari,T., Yau,K.-W. and Nathans,J.
TITLE Direct Submission
Submitted (02-JAN-2004) Molecular Biology, Neuroscience, Johns
Hopkins University, 725 N. Wolfe Street, Baltimore, MD 21205, USA

FEATURES
source
1..1758
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..1758
/note="chloride channel"
/codon_start=1
/product="beactrophin 1"
/protein_id="AA89654.1"
/db_xref="GI:41216873"
/translation="MTITTSQVANNALGSFSLLCWRGSIYKLYGFIPLICY
IIRTYRLALEEQLMFEKLTLYCDYIQILPISFVLSFYVTVVTRMNQENLFW
PDRMLSVGFVEGKDEQRLRRTLLRYANLGNVLILRSVSTAIVYRFPASQHLVOA
GFWTPAEHKQLEKSLPHNMFWVWVFMILSMKAMLGRIIDPILLOSILNMTLR
TCGHLAYADWISIPLYTOVVYVAVSFLTCVGRQFLNPKAKAYGHELDLVVPVF
TPIQFFPYGMILKVAEQLINPGEDEDDPFTNMIVDNRLQVSLAAYDEMHDLPMBP
DMWNKRPQPTTAAAOCPRRASFMGSTNINISLNKEMEFQRODEBEBAHAGIICR
FLGQSDHHPPRANSTRKILMPRESLHLEGPKNHAKAKONVQGEENKAMKLVKAV
DARSAFLYORPGYSAPQPLSPTPMFPPLPSAPBKASHVGIIDKDSLKTIVSSG
AKKSFELISDEGALMEHPVSQVRRKTFEFLNLDMEIPENHLKEPLEOSPNIHTT
LKDHMDPYMALENRDEAHS"

ORIGIN

Alignment Scores:

Pred. No.: 1.68e-266 Length: 1758
Score: 3110.00 Matches: 583
Percent Similarity: 99.66% Conservat: 0
Best Local Similarity: 99.66% Mismatches: 2
Query Match: 99.68% Indels: 0
Gaps: 0

US-09-622-964A-3 (1-585) x AY515704 (1-1758)

QY 1 MetThrIleThrTrpThrSerGlnValAlaAsnAlaArgLeuGlySerPheSerArgLeu 20
Db 1 ATGACCTTCACTTACACCAAGCAAGTGGCTTAATGCCGCTTAAGAGCTCTTCTCCCGCTG 60
QY 21 LeuLeuCysTrpArgTrpArgGlySerIleTrpIleLeuLeuLeuLeuLeuLeuLeu 40

Db 61 CTGCTGTGCTGGCGGGGACACATCTACAGCTGCTATATGACGAGTCTTAATCTTCCTG 120

QY 41 LeuCysTrpTrpIleIleIleArgPheIleTrpArgLeuAlaLeuThrGluGluGluGluLeu 60
Db 121 CTCTGCTTACCTTACCAACCCGCTTATTTATTAAGCTGGCCCTTCACGGAAGAACACACTG 180

QY 61 MetPheGluLeuLeuThrLeuLeuTrpCysArgSerTrpIleGlnLeuIlePhePheIleSerPhe 80
Db 181 ATGTTTAGAAATGACTCTGTATTTGGACAGACTACATCCAGCTCATCCCATTTCTCTTC 240

QY 81 ValLeuGlyPheTrpValIleThrLeuValIleThrArgTrpTrpAsnGlnTrpGluAsnLeu 100
Db 241 GTGCTGGGCTTTACAGTACGCTGCTGTGTGTAACCGCTGTGGAAACAGTACGAAACCTG 300

QY 101 ProTrpProAspArgLeuMetSerLeuValSerGlyPheValGluGlyValSerGluGln 120
Db 301 CCGTGGCCGACCGCCCTCATGAGCCCTGTGTGGGCTTCGTGAAAGCAAGACGAGCA 360

QY 121 SerArgLeuLeuArgArgThrLeuIleArgTrpAlaAsnLeuGlyAsnValLeuIleLeu 140
Db 361 GCGCGGCTGCTGGCGGCGACGCTCATCCGCTTACGCACTGGGCAACGTCCTCATCTG 420

QY 141 ArgSerValSerThrAlaValTrpIleArgPheProSerAlaGlnIleLeuValGlnAla 160
Db 421 CGCAGCGTCAAGACCGCAGTCTCAAGCGCTTCCCGCCGCGCCAGCACTGTGTGCAAGCA 480

QY 161 GlyPheMetThrProAlaGluIleIleGlyGlnLeuGluLeuLeuSerLeuProIleAsnMet 180
Db 481 GGCCTTATATGACTCCGGGAAACACAGCAATTTGAGAAATGAGCCCTTACACACACATG 540

QY 181 PheTrpValProTrpValTrpPheAlaAsnLeuSerMetLeuAlaTrpLeuGlyArg 200
Db 541 TTCTGGGTGCTTGGGT 600

QY 201 IleArgAspProIleLeuLeuGlnSerLeuLeuAsnGluMetAsnThrLeuArgTrpGln 220
Db 601 ATCCGGGACCTTATCCGCTGCCAGAGCTGTCTGAAGATGAACACTTGGCTGCTGAC 660

QY 221 CysGlyIleIleLeuTrpAlaTrpAspTrpIleSerIleProLeuValTrpThrGlnVal 240
Db 661 TGTGGACACTGTATGCTTACGCTGATGATGATGATGATGATGATGATGATGATGATGATG 720

QY 241 ThrValAlaValTrpSerPhePheLeuThrCysLeuValGlyArgGlnPheLeuAsnPro 260
Db 721 ACTGTGGCGGTGATGAGCTTCTTCTGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780

QY 261 AlaLeuAlaTrpProGlyIleGluLeuAspLeuValIleProValPheThrPheLeuGln 280
Db 781 GCCAAGGCTTACCTGCGCATGAGCTGACCTGTTGTGCGCTTACAGTTCCTGCAAG 840

QY 281 PhePhePheTrpAlaGlyTrpLeuValAlaGluGlnLeuIleAsnProPheGlyGln 300
Db 841 TTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 900

QY 301 AspAspAspPheGluThrAsnTrpIleValIleAspArgAsnLeuGlnValSerLeuLeu 320
Db 901 GATGATGATGATTTTATGACCAACTGATTTGTGACAGAAATTTGAGAGTGTCTCTGTTG 960

QY 321 AlaValAspGluMetIleGlnAspLeuProArgMetGluProAspMetTrpTrpAsnLeu 340
Db 961 GCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020

QY 341 ProGluProGlnProProTrpTrpAlaIleAsnAlaGlnPheArgArgAlaSerPheMet 360
Db 1021 CCGGAGCCACAGCCCGCTTACAGCTGCTTCCGCGCAAGTTCGCTGAGGCTTCTTTATG 1080

QY 361 GluSerThrPheAsnIleSerLeuAsnGluGluMetGluPheGlnProAsnGlnGlu 380
Db 1081 GAGTTCACCTTCAACATCAGCTTGAACCAAGAGAGATGAGTTCAGGCCCAATCAGAG 1140

QY 381 AspGluGluAspAlaIleAlaGlyIleIleGlyArgPheLeuGluGlnSerTrpAsp 400

Db 1141 GACGAGAGATGCTCAAGCTGCGATCATTTGCGCGCTTCTTCAAGCGCTGCAAGTCCCATGAT 1200
Qy 401 HishisProProArgAlaAsnSerArgThrLysLeuLeuTrpProLysArgLysSerLeu 420
Db 1201 CACCATCTCCACAGGCAAACTCAAGACCAAACTACTGTGGCCCAAGAGAGAAATCCCTT 1260
Qy 421 LeuHsaglLysLeuProLysAsnHsLysAlaAlaLysGlnAsnValArgLysGlnLys 440
Db 1261 CTCACAGAGGCGCTGCCCAAAAACACAAAGGACGCAAAAGATTAAAGGGCCAGGAA 1320
Qy 441 ArgAsnLysAlaTrpLysLeuLysAlaValAspAlaPheLysSerGlyProLeuTyrGln 460
Db 1321 GACAAACAAGGCGCTGGAAGCTTGAAGCTGTGACGCGCTTCAAGTCTGCCCTCACTGATCAG 1380
Qy 461 ArgProGlyTyrTyrSerAlaProGlnThrProLysSerProThrProMetPhePhePro 480
Db 1381 AGGCGAGGCTACTCAAGTGCACACAGAGCGCCCTCAAGCCCACTCCCATGTTCTTCCCC 1440
Qy 481 LeuGluProSerAlaProSerLysLeuHsLysSerValThrGlyLeuAspThrLysAspLys 500
Db 1441 CTAAACAATCATGCGCCCTCAAGCTTCAAGTCTCAAGGCAATGACACAAAGACAA 1500
Qy 501 SerLeuLysThrValSerSerGlyAlaLysLysSerPheGluLeuLysSerGluSerAsp 520
Db 1501 AGCTTAAAGACTGTGAGTTCTGGGGCCAAAGAAAGTTTGAATTGCTCTCAGAGGCGAT 1560
Qy 521 GlyAlaLeuMetGluHsLeuProGluValSerGlnValArgValArgLysThrValGluPheAsn 540
Db 1561 GGGGCGCTTGAAGACACCCAGAAAGTATCTCAAGTGAAGAGAAAGAACTGTGAGTTTAC 1620
Qy 541 LeuThrAspMetProGluLeuProGluLysAsnHsLysLysGluProLysGluGlnSerPro 560
Db 1621 CTGACGGAATACCGAGATCCCGAAATATCACTCAAAAGAACTTTGAAACATCATCA 1680
Qy 561 ThrAsnLeuHsLysThrLeuLysAspHsMetAspProTyrTrpAlaLeuGluAsnArg 580
Db 1681 ACCAACTACACACTACACTCAAGATCATCATGATCTTATTTGGGCTTGGAAACACAG 1740
Qy 581 AspGluAlaHsSer 585
Db 1741 GATGAGCAGCATCC 1755

RESULT 4
AF073501 2171 bp mRNA linear PRI 02-SEP-1998
LOCUS Homo sapiens vitelliform macular dystrophy protein (VMD2) mRNA,
complete cds.
AF073501.1 GI:3511241

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
SOURCE

Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 2171)
Stohr, H., Marguardt, A., Rivera, A., Cooper, P.R., Nowak, N.J.,
Shows, T.B., Gerhard, D.S. and Weber, B.H.
A gene map of the Best's vitelliform macular dystrophy region in
chromosome 11q12-q13.1
Genome Res. 8 (1), 48-56 (1998)
9445487
2 (bases 1 to 2171)
Marguardt, A., Stohr, H., Passemore, L., Kraemer, F., Rivera, A. and
Weber, B.H.P.
Direct Submision
Submitted (23-JUN-1998) Human Genetics, University, Biotzentrum, Am
Hubland, Muerzberg 97074, Germany
Location/Qualifiers
1..2171
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

gene
CDS
/chromosome="11"
/map="11q12-q13.1"
/feature type="retinal pigment epithelium"
1..2171
/gene="VMD2"
89..1846
/gene="VMD2"
/note="Best disease"
/product="vitelliform macular dystrophy protein"
/protein_id="AAC33766.1"
/db_xref="GI:3511242"
/translation="MTITYSQVANAELGFSRLLLCGRSIYKLILGEFLIFLLCY
IIRFTYRLALTEEQLMFEKLTLYCDSYIOLIPISFVLGFYTLVVRMWNQYENLPM
PDRMSLVSGFVEGDEQRLRRITLIRANLGVILRSVTRVAVRPSAOHLVOA
GPMTPARHKLKLSLPHMFWPMWRANLSMKVAGSRIRDLTILLOSILBNMTLR
TQCHLYAYHMISLPIVYTOVTVAYVSFFLTCLVGGQPLNPAKAYRGHELDLVYV
TFIQFFPYGMWKVAEQLINPGEEDDFFTNMIVDSNLQVSLAVDEMHODLPRMB
DMWNKPEQPPYTAASAOFRRASFMGSTENISLNKEMEFQNPQDEEDNAAGIIGR
FLGIQSHDHHPPRANSRKLMPRESLHEGLPKNHKAQKQVRCQEDNKAWMLKAY
DAFSAFLYORPGYSAPOPLSPTPMPFLEPSPLEPSKLSHVTGIDTRKSLKTVSSG
AKSFELSSDGLMEHREPVSOVRRTVEFNLTDMPEIENHLKEFLDSPTIHHT
LKDMDPYMALENDENAS"

ORIGIN

Alignment Scores:

Pred. No.:	2,22e-266	Length:	2171
Score:	3110.00	Matches:	583
Percent Similarity:	99.66%	Conservative:	0
Best Local Similarity:	99.66%	Mismatches:	2
Query Match:	99.68%	Indels:	0
DB:	9	Gaps:	0

US-09-622-964a-3 (1-585) x AF073501 (1-2171)

Qy 1 MetThrIleThrTyrThrSerGlnValAlaAsnAlaArgLeuGlySerPheSerArgLeu 20
Db 89 ATGACATCACTTACACAAAGCAAGTGGCTAATCCCGCTTAAGGCTCTTCCCGCCTG 148
Qy 21 LeuLeuGlySerPheArgGlySerIleTyrLysLeuLeuTyrGlyGluPheLeuIlePheLeu 40
Db 149 CTGCTGGCTGGCGGGGACAGCATCTCAAGCTGCTAATATGGCGAGTCTTAACTTCTCTG 208
Qy 41 LeuCystrTyrTyrIleIleArgPheIleTyrArgLeuAlaLeuThrGluGlnGlnLeu 60
Db 209 CTCTGCTACTACATCAACCGCTTATTTATTTAGCTGGCCCTCAAGGAAGAACACAGCTG 268
Qy 61 MetPheGluLysLeuThrLeuTyrCyAspSerTyrIleGlnLeuIleProIleSerPhe 80
Db 269 ATGTTGAGAAACGACCTGTATTTGGGACAGCTAACATCCAGCTCATCCCATTTCTTTC 328
Qy 81 ValLeuGlyPheTyrValThrLeuValValThrArgTrpTrpAsnGlnTyrGluAsnLeu 100
Db 329 GTGCTGGGCTTCTTACGAGACCTGTGTCGCCCTGGTGGGAACCAAGTACGAAACCTG 388
Qy 101 ProTrpProAspArgLeuMetSerLeuValSerGlyPheValGluGlyLysAspGluGln 120
Db 389 CCGTGGCCGACCCCTCATGAGCTGTGTGGAGCTTCGTCAAGGCAAGAGCAAGCA 448
Qy 121 SerArgLeuLeuArgArgThrLeuIleArgTyrAlaAsnLeuGlyAsnValLeuIleLeu 140
Db 449 GGGCGGCGTGGGGGCGGCGCATCGCTACGCCCAACCTGGGCAACGTGCTCATCTCG 508
Qy 141 ArgSerValSerThrAlaValTyrLysArgPheProSerAlaGlnHsLysLeuAla 160
Db 509 GCGACGCTCAGCAGCGAGCTTACAAAGCGCTTCCCAAGCGCCGACCTGTGTCCAAAGCA 568
Qy 161 GlyPheMetThrProAlaGluHsLysGlnLeuGluLysLeuSerLeuProHsAspMet 180
Db 569 GGCCTTATGACTCCGGCAACACAGCATTTGAGAAACTGACACTTACCAACAACATG 628
Qy 181 PheTrpValProTrpValTrpPheAlaAsnLeuSerMetLysAlaTrpLeuGlyLysArg 200

Db 629 TTCTGGGTGCTGGGTGGTGGTTTGCACACCTGTCATGTAAGCGCTGGCTTGAGAGTGA 688
 Qy 201 ILehrAspProIleuLeuGlnSerLeuLeuEngIUmEAspThrLeuArgThrGln 220
 Db 689 ATCCGGAGCCCTATCTGCTCCAGAGCCTGGTGAACGAGTAACACCTTGAGTACTAG 748
 Qy 221 CysGlyHisLeuYrAlaTyrAspTrpIleSerIleProLeuValTyrThrGlnValVal 240
 Db 749 TGTGGACACCTGTATGCTACGACGATGATAGTATCCACAGTGTATACACAGGTGGTG 808
 Qy 241 ThrValAlaValTyrSerPhePheLeuThrCysLeuValGlyArgGlnPheLeuAspPro 260
 Db 809 ACTGGGGGGGTGACGCTTCTCTGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 868
 Qy 261 AlaValAlaTyrProGlyHisGlnLeuAspLeuValProValPheThrPheLeuGln 280
 Db 869 GCCAAGGCTTACCTGGCCATGAGCTGAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 928
 Qy 281 PhePhePheTyrValGlyTyrLeuValAlaGlnGlnLeuLeuAspProPheGlyGln 300
 Db 929 TTCTTCTTCTATGTTGGCTGGCTGAGGTGGCAGAGCAGCTCATCAACCTTTGGAGAG 988
 Qy 301 AspAspAspAspPheGlnThrAsnTrpIleValAspArgAsnLeuGlnValSerLeu 320
 Db 989 GATGATGATGATTTTGAAGACCACTGAGATTGTCGACAGAAATTTGCAAGGTGCTCCCTG 1048
 Qy 321 AlaValAspGluMetHisGlnAspLeuProArgMetGluProAspMetTyrTrpAsnLys 340
 Db 1049 GCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1108
 Qy 341 ProGluProGlnProProTyrThrAlaAlaSerAlaGlnPheArgArgAlaSerPheMet 360
 Db 1109 CCCGAGCAGCAGCCCTTACACAGCTGCTCCGCCAGCTCCGTCAGAGCTCCCTTATG 1168
 Qy 361 GlySerThrPheAsnIleSerLeuAsnLysGlnGlnMetGluPheGlnProAsnGlnGln 380
 Db 1169 GGCCTCCACCTTCAATCAGCTGACCAAGAGAGATGAGTCCAGCCCAATCAGAGAG 1228
 Qy 381 AspGluGluAspAlaHisAlaGlyIleIleGlyArgPheLeuGlnLeuSerHisAsp 400
 Db 1229 GACGAGAGAGATGCTCAGCTGAGCATTTGCGCGCTTCTTAGGCTGAGCATCCCATGAT 1288
 Qy 401 HisHisProProArgAlaAsnSerArgThrLysLeuLeuTyrProLysArgLysSerLeu 420
 Db 1289 CACCATCTCCCAAGGCAAACTCAAGGACCAAACTAGTGTGCCCAAGAGGAAATCCCTT 1348
 Qy 421 LeuHisGlnGlyLeuProLysAsnHisLysAlaAlaLysGlnAsnValArgGlyGlnGln 440
 Db 1349 CTCACAGAGGCTGCTCCCAAAACCAAGGACCAAAAGCTTAAAGGAGGAGGAGGAG 1408
 Qy 441 AspAsnLysAlaTyrLysLeuValAlaAspAlaPheLysSerGlyProLeuTyrGln 460
 Db 1409 GACAAACAGGCTGGAAGCTTAAAGCTGTGACCCCTTCAAGTGTGCCCACTTATCAG 1468
 Qy 461 ArgProGlyTyrTyrSerAlaProGlnThrProLeuSerProThrPrometPhePhePro 480
 Db 1469 AGGCAAGGCTACTACAGTCCCAAGACACCTTCAAGCCCACTCCCATGTTCTTCCC 1528
 Qy 481 LeuGluProSerAlaProSerLysLeuHisSerValThrGlyIleAspThrLysAspLys 500
 Db 1529 CTAAACACATCAGGCGCTCAAAAGCTTCAAGTCTCAAGGCAATGACCAACCAAGACAAA 1588
 Qy 501 SerLeuLysThrValSerSerGlyAlaLysLysSerPheGluLeuLeuSerGlyLysAsp 520
 Db 1589 AGCTTAAGACTGTGAAGTCTGGGGCCAAAGAAAGTTTGAATGCTCTTCAAGAGAGGAT 1648
 Qy 521 GlyAlaLeuMetGluHisProGluValSerGlnValArgArgLysThrValGluPheAsn 540
 Db 1649 GGGGAGCTTGTAGGACACCAAGAGTATCTCAAGTGAAGAGAAACCTGTGGAGTTTAAAC 1708
 Qy 541 LeuThrAspMetProGluLysProGluAsnHisLysLeuGluProLeuGluGlnSerPro 560
 Db 1709 CTGACGATATGATGAGAGATCCCGAAATCACTCAAGAAAGCTTTGGAACAATCACCA 1768

Qy 561 ThrAsnIleHisThrThrLeuLysAspHisMetAspProTyrTrpAlaLeuGluAsnArg 580
 Db 1769 ACCAACAATACACTACACTCAAGATGATGATGATGATGATGATGATGATGATGATGAT 1828
 Qy 581 AspGluAlaHisSer 585
 Db 1829 GATGAGACACATTC 1843
 RESULT 5
 AF057169
 LOCUS
 DEFINITION
 Homo sapiens bestrorhin (VMD2) mRNA, alternatively spliced product, complete cds.
 AF057169
 AF057169.1 GI:3335158
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS
 1 (bases 1 to 2210)
 Petrukhin, K., Kolet, M.J., Bakall, B., Li, W., Xie, G., Markwell, T., Sandgren, O., Forsman, K., Holmgren, G., Andreasson, S., Vujic, M., Bergen, A.A., McGarvey-Dugan, V., Figueroa, D., Austin, C.P., Metzger, M.L., Caskey, C.T. and Madellius, C.
 Identification of the gene responsible for Best macular dystrophy Nat. Genet. 19 (3), 241-247 (1998)
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 2 (bases 1 to 2210)
 Petrukhin, K.
 REFERENCE
 AUTHORS
 Direct Submission
 Submitted (03-APR-1998) Human Genetics, Merck Research Laboratories, West Point, PA 19486, USA
 JOURNAL
 TITLE
 FEATUERS
 source
 1. .2210
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="11"
 /map="11q13"
 1. .2210
 /gene="VMD2"
 105. .1862
 /gene="VMD2"
 /note="Best's macular dystrophy gene; BMD; alternatively splice isoform"
 /product="bestorhin"
 /protein_id="AA064343.1"
 /db_xref="GI:3335159"
 /translation="MTTITTSQVANNRLCSFRLICWFGSIYKLLYGEPLIFLCY
 IIRFIRLALTEBQQLMEFKTLTCDYIQLPISVIGFVYTLVVTWNNQENLPM
 PRLMLVSGFVGEKDEQRLRLTLIRVANIENVILRSVTAAYKRPPSQHLOA
 GEWTPEHKLQELSLPHNMFVPMVAFNLISMKALGRIIDPILLOSLINEMTLR
 TOGCHYAYDWISIPLYOVTVAYSFPLICVGRLOINPAKAYGHELDLVVVF
 TPLQEPFVYGVKVAEOLINPGEDDDPDEMTWIDRNLOVGLIADMEHODLPREV
 DMVWKPEPQPYTASAOFRASFMGTFNLSLKENMEFQPNODEDEDAIIGIGR
 FLGLSHDHRPPANSRTYLLMPKRESLHBLPKNHKAKQNVQRQEDNKAKWKAV
 DAFKSLDHPRQVSAQPTPLSPFMFPLPFSASKLSYTGIDTRKRSKTVSSG
 AKKSFLLSESGALMEHPEVGSQVRKTVFNLTPMPEIPENHLKEPLBQSPNIIHT
 LNDHMDPYALNREDAHS"
 ORIGIN
 Alignment Scores:
 Pred. No.: 2, 28e-266 Length: 2210
 Score: 3110.00 Matches: 583
 Percent Similarity: 99.66% Conservative: 0
 Best Local Similarity: 99.66% Mismatches: 2
 Query Match: 99.68% Indels: 0
 DB: 9 Gaps: 0

US-09-622-964A-3 (1-585) x AF057169 (1-2210)

Qy 1 MetThrIleThrTyThrSerGlnValAlaAsnAlaArgLeuGlySerPheSerArgLeu 20
 Db 105 ATGACCATCACTTACACAGCCAGTGGCTAATGCCGCTTAGGCTCTTCTCCGCGCTG 164
 Qy 21 LeuLeuCystrPArgIySerIleTyTylsLeuLeuTyrgIyGlnPheLeuIlePheLeu 40
 Db 165 CTGCTGATGCGCGGAGCATCTACAGCTGCTAATAGCGAGTTCTTAATCTTCTCGT 224
 Qy 41 LeuCystrTyTylIleIleArgPheIleTyTyrArgLeuAlaLeuThrGlnGlnGlnLeu 60
 Db 225 CTCTGCACATCAATCATCCGCTTAATTAATGAGTGGCCCTCAGAGAAACAACACCTG 284
 Qy 61 MetPheGlnTyLeuThrLeuTyTyrCyseAspSerTyTylIleGlnLeuIleProIleSerPhe 80
 Db 285 ATGTTTGAAGAACTGACTCTTATTTGGACAGCTACATCCAGCTCATCTCCATTTCTTC 344
 Qy 81 ValLeuGlyPheTyTyrValThrLeuValAlaThrArgTyrTyrAsnGlnTyTylsAsnLeu 100
 Db 345 GTGCTGGGCTTCTAAGTAGCGCTGCTGCTGAGACCGCTGCTGGAACCAAGTACGAGACCTG 404
 Qy 101 ProTyrProAspArgLeuMetSerLeuValSerGlyPheValGlnGlyLysAspGlnGln 120
 Db 405 CCGTGGCCCGACCCGCTCATGAGCTGGTGTCTGGGCTTCGTGGAAGGCAAGAGAGCAA 464
 Qy 121 SerArgLeuLeuArgArgThrLeuIleArgTyrAlaAsnLeuGlyAsnValIleuIleLeu 140
 Db 465 GGGCGGCTGCGCGGCGACGCTCATCCGTAACCACTGGGCGCAAGCACTCATATCTCG 524
 Qy 141 ArgSerValSerThrAlaValAlaTyTylsArgPheProSerAlaGlnIleLeuValGlnAla 160
 Db 525 CGCAGCCTCAGCACCGAGCTTACAGCGCTTCCCGACGCGCCAGCACCTGTCAGAGCA 584
 Qy 161 GlyPheMetThrProAlaGlnIleGlyGlnLeuGlyLysSerLeuProIleAsnMet 180
 Db 585 GGGCTTATGACTCCGGCAGAACACAGAGATGGAGAACTGAGACCTTACACACAACTAG 644
 Qy 181 PheThrValProTyrValThrPheAlaAsnLeuSerMetLysAlaTyrLysGlyArg 200
 Db 645 TTCTGGGCGCTGCGGTGGTGGTTCACACTGTCAAGAGGGGTGGCTGGAGGTGCA 704
 Qy 201 IleArgAspProIleLeuLeuGlnSerLeuLeuAsnGlnMetAsnThrLeuArgThrGln 220
 Db 705 ATCCGGAACCTTATCCGCTCAGAGCTGCTGAGACAGATGAACACTTGCATATCTCAG 764
 Qy 221 CysGlyIleLeuTyTyrAlaTyTyrAspTyrIleSerIleProLeuValTyTyrGlnValAla 240
 Db 765 TGTGGAACACCTGTATGCTTACAGCTGATTAATCCACTGGTGTATACACAGGTGGTG 824
 Qy 241 ThrValAlaValTyTyrSerPhePheLeuThrCysLeuValGlyArgGlnPheLeuAsnPro 260
 Db 825 ACTGTGGCGGTGACAGCTTCTCTGACTGTGCTAGTTGGCGGCAAGTCTCTCAACCCA 884
 Qy 261 AlaIysAlaTyTyrProGlyIleGlnLeuAspLeuValAlaProValPheThrPheLeuGln 280
 Db 885 GCCAAGGCTTACCTGGCCATGAGCTGACCTCGTTGGCCCGCTTTCACCTTCTGAG 944
 Qy 281 PhePhePheTyTyrValGlyTyrPheLeuValAlaGlnGlnLeuIleAsnProPheGlyGln 300
 Db 945 TTCTTCTTCTATGTTGGCTGCTGAGAGTGGCAGAGCAAGCTCATCAACCCCTTTGGAGAG 1004
 Qy 301 AspAspAspAspPheGlnThrAsnTyrIleValAspArgAsnLeuGlnValSerLeuLeu 320
 Db 1005 GATCATGATGATTTTGGACCAACTGGAATGTCCGACAGGAATTTGCAAGGTCTCCCTGTG 1064
 Qy 321 AlaValAspGlnMetHisGlnAspLeuProArgMetGlnProAspMetTyTyrTyrAsnLys 340
 Db 1065 GCTGTGATGATGATGACACAGCACTGCTGATGAGCGGAGCATGTACTGGAATAG 1124
 Qy 341 ProGlnProGlnProTyrTyrThrAlaIleSerAlaGlnPheArgArgAlaSerPheMet 360
 Db 1125 CCGAGGCAAGCCCGCTTACACAGCTGCTCCGCGCAAGTTCCGTGAGCACTCCCTTATAG 1184

Qy 361 GlySerThrPheAsnIleSerLeuAsnLysGlnGlnMetGlnPheGlnProAsnGlnGln 380
 Db 1185 GGTCCACCTTCAACATCACTGAAACAAGAGATGAGTTCAGGCCCAATCAAGAG 1244
 Qy 381 AspGlnGlnAspAlaHisAlaGlyIleIleGlyArgPheLeuGlyLeuGlnSerHisAsp 400
 Db 1245 GACGAGAGAGATGCTCAGCTGGCATTTGGCGGCTTCTTCAAGGCTGCAGTCCATGAT 1304
 Qy 401 HisIleProProArgAlaAsnSerArgThrTylsLeuLeuThrProLysArgLysSerLeu 420
 Db 1305 CACCATCTCCCAAGGGCAAACTCAAGSACCAACTTGTGGCCCAAGAGGAATCCCTT 1364
 Qy 421 LeuHisGlnGlyLeuProLysAsnHisLysAlaAlaLysGlnAsnValArgGlyGlnGln 440
 Db 1365 CTCACAGAGGCGCGCCCAAAAACAACAAGCAGCCAAACAAGACGTTAGGGCCAGAA 1424
 Qy 441 AspAsnLysAlaTyrPylsLeuLysValAlaAspAlaPheLysSerGlyProLeuTyTyrGln 460
 Db 1425 GACAAACAAGGCTGGAAGCTTAAAGCTGTGACGCTTCAAGTCTGCCCACTGTATCAG 1484
 Qy 461 ArgProGlyTyTyrTyrSerAlaProGlnThrProLeuSerProThrProMetPhePhePro 480
 Db 1485 AGGCAAGCTTACTACAGTGGCCCAAGAGCGCCCTCAGCCCACTCCATGTCTTCTCC 1544
 Qy 481 LeuGlnProSerAlaProSerLysLeuHisSerValThrGlyIleAspThrLysAspLys 500
 Db 1545 CTAGAACCATCAGGCGCGCTCAAGCTTCAAGTCTCAGAGCATAGACACCAAGACAA 1604
 Qy 501 SerLeuTyThrValSerSerGlyAlaLysLysSerPheGlnLeuLeuSerGlnSerAsp 520
 Db 1605 AGCTTAAAGACTGAGTCTTGGGGCCAAAGAAATTTGAATGTCTCAGAGAGCGAT 1664
 Qy 521 GlyAlaLeuMetGlnHisProGlnValSerGlnValArgArgLysThrValGlnPheAsn 540
 Db 1665 GGGGCTTGTAGGAGCAACCAAGATGTCCTCAAGTGAAGGAAACGTGGAGTTTAAAC 1724
 Qy 541 LeuThrAspMetProGlnIleProGlnLysAsnHisLeuLysGlnProLeuGlnGlnSerPro 560
 Db 1725 CTGACGATATGCGCAGAGATCCCGAATAATCACTCAAAAGACTTTGGAAATTCACA 1784
 Qy 561 ThrAsnIleHisThrThrLeuLysAspHisMetAspProTyrTyrAlaLeuGlnLysAsnArg 580
 Db 1785 ACCAAGATACACTACACTCAACAAAGATCAATGATCTTATTTGGCTTGGAAAAACAGG 1844
 Qy 581 AspGlnAlaHisSer 585
 Db 1845 GATGAAGCACATTC 1859

RESULT 6
 LOCUS CQ729633 2170 bp DNA linear PAT 03-FEB-2004
 DEFINITION Sequence 15567 from Patent WO02068579.
 ACCESSION CQ729633
 VERSION CQ729633.1 GI:42301252
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
 Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
 Kites, such as nucleic acid arrays, comprising a majority of
 humanexons or transcripts, for detecting expression and other uses
 thereof
 Patent: WO 02068579-A 15567 06-SEP-2002;
 JOURNAL PE Corporation (NY) (US)
 FEATRES Location/Qualifiers
 SOURCE 1..2170
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 5,66-264 Length: 2170
 Score: 3083.00 Matches: 582
 Percent Similarity: 99.49% Conservative: 0
 Best Local Similarity: 99.49% Mismatches: 3
 Query Match: 98.81% Indels: 1
 DB: 6 Gaps: 0

US-09-622-964a-3 (1-585) x CQ729633 (1-2170)

QY 1 MetThrIleThrYTrSerGlnValAlaAsnAlaArgLeuGlySerPheSerArgLeu 20
 Db ATGACATCACTTACACAGCCAGG-GCTATATCCCGCTTATAGGCTCTTCTCCGCTG 147
 QY 21 LeuLeuGlyTrpArgGlySerIleTyrlsLeuLeuTyrglyGluPheLeuIlePheLeu 40
 Db CTGCTGAGCTGGCGGGGCGCATCTTACAGCTGCTATATGGCGAGTCTTATCTTCTG 207
 QY 41 LeuGlyTrpYTrIleIleArgPheIleTyrlsArgLeuAlaLeuThrGluGlnGlnLeu 60
 Db CTCTGCTACTACATCATCCGCTTATTTATAGCTGGCCCTCAGGAAGAACACAGCTG 267
 QY 61 MetPheGluLysLeuThrLeuTyrlsCysAspSerTyrlleGlnLeuIleProIleSerPhe 80
 Db ATGTTTGAAGAACTGACTGTGTATGGCAGACATCCAGCTCATCCCATTTCTCTTC 327
 QY 81 ValLeuGlyPheTyrlsValIleThrLeuValIleThrArgTrpTrpAsnGlnTyrglyAsnLeu 100
 Db GTGCTGGAGCTTCTAGCGACCTGTGTGTGACCGCGTGGTGAACACATGACAGACCTG 387
 QY 101 ProTrpProAspArgLeuMetSerLeuValSerGlyPheValGluGlyLysAspGluGln 120
 Db CCGTGGCCCAACCGCTCATGAGCTGTGTGTGGCTTCCGCGAAGGCAAGACAGACAA 447
 QY 121 SerArgLeuLeuAlaGlyThrLeuIleArgTyrlsAsnLeuGlyAsnValLeuIleLeu 140
 Db GGCCTGGCTGCGCGCCGACGCTCATCCGCTACCCAGCCGAGCAACCTGGCTCATCTG 507
 QY 141 ArgSerAlaSerThrAlaValIleTyrlsArgPheProSerAlaGlnIleLeuValGlnAla 160
 Db CGAGCGTACGACCGGACGCTTACAGGCGCTTCCCGAGCGCCGACACCTGGTTCAGCA 567
 QY 161 GlyPheMetThrProAlaGlnIleAlaGlnLeuGlyLysLeuSerLeuProIleAsnMet 180
 Db GCGCTTATGACTCCGGCAGAACACACAGCATTTGGAAGAACTGACCTTACACACAACTG 627
 QY 181 PheTrpValProTrpValIlePheAlaAsnLeuSerMetLysAlaTrpLeuGlyValArg 200
 Db TTCTGGGTGCGCTGGGTGTGTGTGCTTCCAACTGTCAATGAAGGTGGCTTGGAGTCA 687
 QY 201 IleArgAspProIleLeuLeuGlnSerLeuLeuAsnGlnMetAsnThrLeuArgThrGln 220
 Db ATCCGGAACCTTATCTGCTTCCAGAGCTGCTGACAGATGAACACCTTGGCTGACTCAG 747
 QY 221 CysGlyHisLeuTyrlsAlaTyrlsAspTrpIleSerIleProLeuValTyrlsGlnValVal 240
 Db TGTGGAACACCTGATGCTAGCTGATTAAGTATCCACCTGCTATACACAGTGTGTG 807
 QY 241 ThrValAlaValIleTyrlsArgPheLeuThrCysLeuValGlyArgGlnPheLeuAsnPro 260
 Db ACTGTGGCGGTGATACACTTCTCTGACTTGTCTAGTTGGGCGCGAGTTTCTGAAACCA 867
 QY 261 AlaLeuAlaValTyrlsProGlyHisGlnLeuAspLeuValIleProValPheThrPheLeuGln 280
 Db GCCAAGGCTTACCTGCGCATGAGCTGACCTCTGTGGCCCGCTTCACTTCTCTCAG 927
 QY 281 PhePhePheTyrlsValGlyTrpLeuLysValAlaGlnGlnLeuIleAsnProPheGlyGlu 300
 Db TTCTTCTTCTATGTTGGCTGGCTGAAGTGGCAGACGCTCATCAACCTTTTGGAGAG 987
 QY 301 AspAspAspAspPheGluThrAsnTrpIleValAspArgAsnLeuGlnValSerLeuLeu 320

Db 988 GATGATGATGATTTTGAACCAACTGGATTGTGACAGGAAATTTGCAGGTCTCCCTGTTG 1047
 QY 321 AlaValAspGluMetHisGlnAspLeuProArgMetGluProAspMetTyrlsPAsnLys 340
 Db GCTGTGTGATGATGATGACACGAGACTGTGCTCGATGAGCGGACGACATGTACTGGAATAG 1107
 QY 341 ProGluProGlnProProTyrlsThrAlaAlaSerAlaGlnPheArgAlaAspPheMet 360
 Db CCCGAGCCACAGCCCCCTTACAGAGCTGTCCGCGCAAGTTCGTCAGAGCTCTCTTATG 1167
 QY 361 GlySerThrPheAsnIleSerLeuAsnLysGluMetGluPheGlnProAsnGlnGlu 380
 Db GGCTCCACCTTCAACATCAGCTGAAACAAAGAGAGATGAGATTCCAGCCCAATCAGAG 1227
 QY 381 AspGluGluAspAlaHisAlaGlyIleIleGlyArgPheLeuGlyLeuGlnSerHisAsp 400
 Db GACGAGGAGGATGCTCAGCGTGGCATCATTTGGCGCTTCTAGGCGCTGACGTCCATGAT 1287
 QY 401 HisHisProProAlaAlaAsnSerArgThrLysLeuLeuTrpLysValArgGluSerLeu 420
 Db CACCATCTCCAGGCGCAACTCAGAGCCAACTACTGTGGCCCAAGAGGAAATCCCTT 1347
 QY 421 LeuHisGlnGlyLeuProLysAsnHisLysAlaAlaLysGlnAsnValArgGlyGlnGlu 440
 Db CTCACAGAGGCTGCGCCCAAAACCAACAGGACGCCAAACAGACGTTAGGGCCAGAGA 1407
 QY 441 AspAsnLysAlaTrpLysLeuLysAlaValAspAlaPheLysSerGlyProLeuTyrlsGln 460
 Db GACACACAGGCTCGAAGCTTAAAGCTGTGAGCGCTTCAAGTCTGCGCCACATGTATCAG 1467
 QY 461 ArgProGlyTyrlsSerAlaProGlnThrProLeuSerProThrProMetPhePhePro 480
 Db AGGCGAGGCTACTACAGTGGCCCAACAGACGCCCCCTCAGCCCATCTCCATGTTCTTCCC 1527
 QY 481 LeuGluProSerAlaProSerLysLeuHisSerValThrGlyIleAspThrLysAspLys 500
 Db CTAGAACATACAGCGCGGTAAAGCTTACAGGTGTCAAGGCAATGACACCAAGACAA 1567
 QY 501 SerLeuLysThrValSerSerGlyAlaLysLysSerPheGluLeuLeuSerGluSerAsp 520
 Db AGCTTAAAGCTGTAGTCTGTGGGCGCAAGAAAGTTTGAATTCCTCAGAGAGCAT 1647
 QY 521 GlyAlaLeuMetGluHisProGluValSerGlnValArgArgLysThrValGluPheAsn 540
 Db GGGGCTTGTAGAGCACCCAGAAATATCTCAAGTGGAGGAAACCTGGAGGTTTAACT 1707
 QY 541 LeuThrAspMetProGluIleProGluAsnHisLysLysGluProLeuGluGlnSerPro 560
 Db CTAGCGATATGCGACAGATCCCGAAATACCTCAAGAACCTTTGGAACTATCACA 1767
 QY 561 ThrAsnIleHisThrThrLeuLysAspHisMetAspProTyrlsAlaLeuGluAsnArg 580
 Db ACCAAGCTTACACCTACACTCAAAAGATCATGATCTTATTTGGGCTTGGAAAAACAGG 1827
 QY 581 AspGluAlaHisSer 585
 Db 1828 GATGAAGCACATTC 1842

RESULT 7
 BD136721 2429 bp DNA linear PAT 18-SEP-2002
 LOCUS BD136721
 DEFINITION Best's macular dystrophy gene.
 ACCESSION BD136721
 VERSION BD136721.1 GI:23231666
 KEYWORDS JP 2002504559-A/3.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 2429)
 AUTHORS Petrukhin,K., Caskey,T.C., Metzker,M. and Wadelius,C.
 TITLE Best's macular dystrophy gene
 JOURNAL Patent: JP 2002504559-A 3 12-FEB-2002;

Db 1784 ACAGCATAGACACCAAGACAAAGCCTTAAGACTGTGATGTTCTGGGCCCAAGAAAGT 1843

QY 513 PheGluLeuLeuSerGluSerAspGluValLeuMetGluHisProGluValSerGluVal 532

Db 1844 TTTGAATTGCTCTCAGAGAGCGATGGGCGCTTGATGAGACCCCAAGATATCTCAAGTG 1903

QY 533 ArgArgLeuThrValGluPheLeuLeuThrAspMetProGluLeuProGluAsnHisLeu 552

Db 1904 AGAGAGAAACTGTGGAGTTTAACCTGACGGATATGCCAGAGATCCCGAAATATCACTCC 1963

QY 553 LysGluProLeuGluGlnSerProThrAsnLeuHisThrThrLeuLeuAspHisMetAsp 572

Db 1964 AAAAGAACCTTTGGAAATACCAACCAACATACATACATCACTCAAGATCACAATGAT 2023

QY 573 ProTyrTrpAlaLeuGluAsnArgAspGluAlaHisSer 585

Db 2024 CCTTATGGGCGCTTGGAAAAAGGAGTGAAGCACTTCC 2062

RESULT 8

LOCUS AF057170 2420 bp mRNA linear PRI 17-OCT-1998

DEFINITION Homo sapiens bestrophin (VMD2) mRNA, alternatively spliced product, complete cds.

ACCESSION AF057170.1 GI:3335160

VERSION AF057170.1 GI:3335160

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo. 1 (bases 1 to 2420)

AUTHORS Sandgren, O., Forsman, K., Holmgren, G., Andreasson, S., Vujic, M., Bengtson, A., McGarvey, D., Figueroa, D., Austin, C.P., Metzger, M.L., Caskey, C.T. and Wadelius, C. Identification of the gene responsible for Best macular dystrophy

TITLE Nat. Genet. 19 (3), 241-247 (1998)

JOURNAL 98324772

MEBLINE 9662395

REFERENCE 2 (bases 1 to 2420)

AUTHORS Petrushkin, K.

TITLE Direct Submission

JOURNAL Submitted (03-APR-1998) Human Genetics, Merck Research Laboratories, West Point, PA 19486, USA

FEATURES

source

1..2420

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/chromosome="11"

/map="11q13"

1..2420

/gene="VMD2"

105..1412

/gene="VMD2"

/note="Best's macular dystrophy gene; BMD; alternatively splice isoform; contains exon 7"

/codon_start=1

/product="bestrophin"

/protein_id="AAC54344.1"

/db_xref="GI:3335161"

/translation="MTTITTSQVANAFLGSPSRLLCWRGSIYKLYGEPFLILCY IIRITVALTEEQLMPEKLTLYCDYIQLIPISFLVGFVTVLVTRMNOYENLFW PDRMLSVSGFEKDEQGRILRLTILRYANLGNVLIRSVTAIVRRPFAOHLVOA GPMTPABHQLLEKLSPHNMFWVWVFMKAMLGIRIDPILLOSILNMTTLR TOGHLAYADWISILPIVYTOVVAVVSFLTCLVGRQPLNPBKAYGHELDLVVPP TPQLFFPYVMILVGLSRALLGWRHGRGGOGLSTRMCGQRKSRVSSQAWMT PVLPATREAGSLEBGRRLMWSSSTPLEMMMLILRTGLSTGICRCPCWLMWR CTRICGWSRTGISTPSHSPITQLPPSSVEPLVAPSTSA"

ORIGIN

Alignment Scores:

Pred. No.: 3.17e-260 Length: 2420

Score: 3041.50 Matches: 582

Percent Similarity: 89.13% Conservative: 0

Best Local Similarity: 89.13% Mismatches: 3

Query Match: 97.48% Indels: 69

DB: 9 Gaps: 1

US-09-622-964A-3 (1-585) x AF057170 (1-2420)

QY 1 MetThrIleThrTyrThrSerGlnValAlaAsnAlaArgLeuGlySerPheSerArgLeu 20

Db 105 ATGACATCACTTACACAAAGCAAGTGGCTAATCCCGCTTAAGCTCTCTCCCGCTG 164

QY 21 LeuLeuGlyTrpArgGlySerIleTyrIleLeuLeuTyrGlyGluPheLeuIlePheLeu 40

Db 165 CTGCTGCTGCTGGCGGGCAGATCTTACAGCTGCTATATGGCAGTTCTTATTTCTTG 224

QY 41 LeuGlyTyrIleIleIleArgPheIleTyrArgLeuAlaLeuThrGluGluGlnLeu 60

Db 225 CTGCTACTACTACATCACTCCGCTTATTTATAGCTGGCCCTCACGGAAACACAGCTG 284

QY 61 MetPheGluLysLeuThrLeuTyrCysAspSerTyrIleGlnLeuIleProIleSerPhe 80

Db 285 ATGTTAGAAATGACTCTGTATTTGGACAGCTACATCCAGCTCATCCCATTTCTTC 344

QY 81 ValLeuGlyPheTyrValThrLeuValValThrArgTrpAsnGlnTyrGluAsnLeu 100

Db 345 GTGCTGGCTTCTACAGTACGAGCTGTGTGTGTATCCCGCTGTGGAAACAGTACGAACTG 404

QY 101 ProTrpProAspArgLeuMetSerLeuValSerGlyPheValGluGlyLysAspGluGln 120

Db 405 CCGTGGCCCAACCGCCCTCATGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 464

QY 121 SerArgLeuLeuArgArgThrLeuIleArgTyrAlaAsnLeuValLeuIleLeu 140

Db 465 GCGCGCTGT 524

QY 141 ArgSerValSerThrAlaValTyrLysArgPheProSerArgIleHisLeuValGlnAla 160

Db 525 CCGAGCGTACGACCGGAGTCTACAGCGCTTCCAGGCGCCAGACGCTGTGTGTGTGTGTGT 584

QY 161 GlyPheMetThrProAlaGluHisLysGlnLeuGluLysLeuSerLeuProHisAsnMet 180

Db 585 GGCCTTATGACTCCGCGACAAAGACAGTGTGAAGAACTGAGCTTACACCAACATG 644

QY 181 PheTrpValProTrpValTrpPheAlaAsnLeuSerMetLysAlaThrLeuGlyValArg 200

Db 645 TTCTGTGCTGCTGT 704

QY 201 IleArgAspProIleLeuLeuGlnSerLeuLeuAsnGluMetAsnThrLeuArgThrGln 220

Db 705 ATCCGGGACCCCTATCTGCTCCAGAGCTGTGTGAACGAGATGAACACTTGGGTACTCAG 764

QY 221 CysGlyHisLeuTyrAlaTyrAspTrpIleSerIleProLeuValTyrThrGlnValVal 240

Db 765 TGTGACACCTGTATCTTCAAGCTGATTAATCCACTGGTATACACAGGTGGTG 824

QY 241 ThrValAlaValTyrSerPhePheLeuThrCysLeuValGlyArgGlnPheLeuAsnPro 260

Db 825 ACTGTGGCGGTGACAGCTTCTTCTGCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 884

QY 261 AlaValAlaTyrProGlyHisGluLeuAspLeuValValProValPheThrPheLeuGln 280

Db 885 GCCAAGGCTTACCTGTGGCATGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 944

QY 281 PhePhePheTyrValGlyTyrPheLysVal 290

Db 945 TTCCTTCTATGT 1003

QY 290 290

Db 1004 GCATGGCCAGAGGGGTATGGCCAGACGCTGTTAGACGAGAGATGCACTGTCAAGAAAG 1063

QY 290 290

LOCUS	DEFINITION	ACCESSION	VERSION	2187 bp	mRNA	linear	PRI	26-AUG-2003
Db	1064	GAAGTCTCAGGGGTAAAGCAGCCAGGCGGTGTGGGCGACACCTGTATCCACGTAAC	1123					
Qy	291	-----AAGlu	292					
Db	1124	TCCGAGAGCTGAGCAGAGAAATCGCTTGAACCCGGAGCGGAGAGTTGTGTGGCAGAG	1183					
Qy	293							
Db	1184	CAGCTCATCAACCCCTTTGGAGAGGATATGATGATTTTGGAACCAACTGTGATGTCAGC	1243					
Qy	313	ArgAenLeuGlnValSerLeuLeuAlaValAerGumEthiAglAerLeuProArgMet	332					
Db	1244	AGGAATTTGCAGGTCTCCCTGTGGCTGTGATGAGATGCACAGAGACTGCTCGATG	1303					
Qy	333	GluProAerMetTyTTrpAsnValProGluProGluProGluProTyTTrpAlaAserAla	352					
Db	1304	GAGCCGGACATGTATGTGAATAGGCCGAGCCAGCCGCCCTTACACAGCTGCTTCGCC	1363					
Qy	353	GlnPheAerGluAlaSerPheMetCylSerThrPheAsnIleSerLeuAsnValGluGlu	372					
Db	1364	CAGTTCCTCGAGCTCTCTTATGGGCTCCACTTTCAATCAATCAAGCTTAACAAGAGAG	1423					
Qy	373	MetGluPheGlnProAerGlnGluAerGluGluAerAlaValAglValIleGluArg	392					
Db	1424	ATGAGATTCCAGCCCAATCAGAGAGACAGAGAGATGCTCAGCGTGGCATCATTTGGCCG	1483					
Qy	393	PheLeuGluLeuGlnSerHisAerHisHisProProAerGluAaAerAerGluThrLeuLeu	412					
Db	1484	TTCCTAGGCGCTGCAGCTCCCATGATCAGCATCTCCAGGCGAATCTCAAGACCAACTA	1543					
Qy	413	LeuTrpProValuArgGluSerLeuLeuHisGluGluLeuProValuAerHisGluValAla	432					
Db	1544	CTGTGGCCCAAGAGGAAATCCCTTCTCCACAGAGGCGCTGCCCAAAACCAAGAGCACCC	1603					
Qy	433	LywGlnAerValArgGluGlnGluAerAsnValAerTrpValuLeuValAlaAerAla	452					
Db	1604	AAACGAAACGTTTAGGGCGCCAGGAACACAAAGAGCTTGAAAGCTTAAAGCTGTGGACGCC	1663					
Qy	453	PheLeuSerGluProLeuTyGlnAerProGluTyTyTySerAlaProGluThrProLeu	472					
Db	1664	TTCAAGTCTGCCCCCACTGTATCAGAGCGCAGGCTCTCAGTGCAGCCCAAGACGCCCTTC	1723					
Qy	473	SerProThrProMetPhePheProLeuGluProSerAlaProSerLeuLeuHisSerVal	492					
Db	1724	AGCCCACTCCCATGTTCTTCCCTTAACACCATAGCGCGGTCAAACTTCACAGTGTCTC	1783					
Qy	493	ThrGluIleAerThrLeuAerPheSerLeuTyThrValSerSerGluValuLeuLeuSer	512					
Db	1784	ACAGGCATAGACACCAAAACAAAGACTTAAAGATGTGAAGTCTGGGCGCAAGAAAGT	1843					
Qy	513	PheGluLeuLeuSerGluSerThrAerGluValuLeuMetGluHisProGluValuSerGlnVal	532					
Db	1844	TTTGAATAGTCTCTCAGAGAGCGATGGGGCTTTGATGAGAGCAACCCAGAAAGTATTCACATGTG	1903					
Qy	533	ArgArgLeuThrValGluPheAsnLeuThrAerMetProGluIleProGluAsnHisLeu	552					
Db	1904	AGAGAGAAACCTGTGAAGTTTAACTGACGGATATGCCAGAGATCCCGAAATATCACTCTC	1963					
Qy	553	LywGluProLeuGluGlnSerProThrAsnIleHisThrThrLeuTyAerHisMetAer	572					
Db	1964	AAAGAACTTTGGAAACATACCAACCAACATACATACATCACTAAAGATCAGATGGAT	2023					
Qy	573	ProTyTTrpAlaLeuGluAsnArgAerGluAlaHisSer	585					
Db	2024	CCTTATTTGGGCTTGAAAAACAGGAGTGAAGACATTTCC	2062					

KEYWORDS	Macaca fascicularis (crab-eating macaque)		
SOURCE	Macaca fascicularis		
ORGANISM	Eukaryota; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Macaca.		
REFERENCE	1 (bases 1 to 2187) Okamoto,H., Umeda,S., Suzuki,M.T., Yoshikawa,Y., Tanaka,Y. and Iwata,T. Molecular Cloning of VMD2 Gene from Cynomolgus Monkey (Macaca fascicularis) Unpublished 2 (bases 1 to 2187) Okamoto,H., Umeda,S., Suzuki,M.T., Yoshikawa,Y., Tanaka,Y. and Iwata,T. Direct Submission Submitted (01-AUG-2003) National Tokyo Medical Center, National Institute of Sensory Organs, 2-5-1, Higashtagaoka, Meguro-ku, Tokyo 152-0021, Japan		
JOURNAL	Location/Qualifiers		
REFERENCE	1. .2187 /organism="Macaca fascicularis" /mol_type="mRNA" /db_xref="taxon:9541" 1. .2187 /gene="VMD2" 67. .1824 /gene="VMD2" /codon_start=1 /product="destrorhin" /protein_id="AA056049.1" /db_xref="GI:34013783" /translation="MITTITTSQVNNARLGSFRRLLLCMRGSIYKLYGEPFIFLLCY IRIRYRLALBEOQLMEFKLTICDSYIOLIPSPFGLPVTLVVRMNOYENLPM PDRMLTSGVSGFVEKDEQGLRPETLRVNLGNVLLRGVSTAVYRPSAOLNOA GEMTPASHKQLEKSLPHNNFWPWPVPAULSKMAGIGIRDPILLOSILNENRTR TQCHLAAYDMISILPLYTVVTVAVSPFLTLCVGRQPLNPAKAYGHELDLVVPE TFLOEPFYVWMLKYAEQLINPFGEDDDPFTNMVLDNLLQVSLAVDEMQDLPRMEP DMVYNEDEPHPPYPAASAOPRASFMGSTFNI SLNKEEMFQPNQDEKDAHTGIR FLGISHDHHPGANSRRTKLMPRESLTHEGLPKHKAVKONVGRGLDNKAMLKAV DAFSAPLVGRPGYYSAPQTPPLSTPMPFPPEBVSVLKLSNYGIDPKDKSLKTVSSG AKNSSELSSGDAKLMEHPEVSHVRKTVENFLDMPEIRFENMLKEPFLSESTIHAT LKHVDVPMALNENDEAHS"		
ORIGIN			
Alignment Scores:			
Pred. No.:	1.11e-256	Length:	2187
Score:	3001.00	Matches:	564
Percent Similarity:	96.92%	Conservative:	3
Best Local Similarity:	96.41%	Mismatches:	18
Query Match:	96.19%	Indels:	0
DB:	9	Gaps:	0
US-09-622-964A-3 (1-585) x AY357925 (1-2187)			
Oy	1 MetThrlleThrYTrnSerGlnValAlaAsnAlaArgLeuGlySerPheSerArgLeu 20		
Db	67 ATGACCATCATTACAGACCAAGTCAGTCAATCCCGCTTACGGCTCTCCGGCTG 126		
Oy	21 LeuLeuCyTrpArgGlySerIleTyrlYsLeuLeuTyrglyGluPheLeuIlePheLeu 40		
Db	127 CTGCTGTCCTGGCGGGACAGCATCTACAGCTGCTGATGCGAGTCTTCATCTTCTG 186		
Oy	41 LeuCyTrpYTrlIleIleArgPheIleTyrlArgLeuAlaLeuThrGluGlnGlnLeu 60		
Db	187 CTCGCTACTACATCATCCGCTTATATTATAGCTGAGCCCTCCACAGAGCAACACTG 246		
Oy	61 MetPheGluYsLeuThrLeuTyrlCysAspSerTyrlIleGlnLeuIlePheIleSerPhe 80		
Db	247 ATGTTTAGAAGCTGACTCTGTATTGGACAGCTACATCCAGCTCATCCCATTTCTTC 306		
Oy	81 ValIleuGlyPheTyrlValThrLeuValThrArgTrpTrpAsnGlnTyrgluAsnLeu 100		

Db 307 GTGCTGGGCTTCTAGCTGACGCTGGTCTGACCCGCGTGGTGGACGAGACGACCTG 366
 Qy 101 ProTppProAspArgLeuMetSerLeuValSerGlyPheValGluGlyLeuAspGluGln 120
 Db 367 CCGTGGCCCGACCGCTCATGAGCTGGTGGGCTTCTGTTGGAAGGACAGACGACGAA 426
 Qy 121 SerArgLeuLeuArgThrLeuIleArgTyrAlaAsnLeuGlyAsnValIleuIleu 140
 Db 427 GGGCGGCTGCTGGCGGACGCTCATCCGCTACGCCAACCTGGGACAGCTGCTATCTTG 486
 Qy 141 ArgSerValSerThrAlaValTyrXylysaArgPheProSerAlaGlnIleuValGlnAla 160
 Db 487 CGCACCGTCAGCACCGCGCTCTACAAAGCGCTTCCCGACGCGCCACACCTGGTGAACA 546
 Qy 161 GlyPheMetThrProAlaGlnHislysglnleuGlnlylyleuSerLeuProHisAsnMet 180
 Db 547 GGCCTTAAGACCCCGGACAGAACAGAGAGTTGGAGAAAGCTGAGTTACACACAAACATG 606
 Qy 181 PheTrrValProTrrValTrrPheAlaAsnLeuSerMetLysAlaTrrPheGlyGlyArg 200
 Db 607 TTTCTGGGTGCCCTGGGTGTGGTTGCCAACCTGTCAATGAAGCGCTGGAGGTGCA 666
 Qy 201 IleArgAspProIleLeuLeuGlnSerLeuLeuAsnGlnMetValSerThrLeuArgThrGln 220
 Db 667 ATCCGGGACCCCTATCCCTGCTCCAGAGCCCTGGAACGAGATGAACACCTTGGCAGCTGAG 726
 Qy 221 CysGlyHisLeuTyrAlaTyrAspTrrIleSerIleProLeuValTyrThrGlnValVal 240
 Db 727 TGTGACACCTGTATGCTTACGACTGAGATGATATCCACGTGGGTATACACAGGTGGTG 786
 Qy 241 ThrValAlaValTyrSerPhePheLeuThrCysLeuValGlyArgGlnPheLeuAsnPro 260
 Db 787 ACTGGGGGGGTGACGCTTCTTCTGACTGTGTAGTTGGCGGAGTTTCTGACCCCA 846
 Qy 261 AlaIysAlaTyrProGlyHisGlnLeuAspLeuValValProValPheThrPheLeuGln 280
 Db 847 GCCAAGGCTACCCCTGAGCATGAGCTGAGCTGCTGTCGCCGCTTCAACATTCCTGAG 906
 Qy 281 PhePhePheTyrValGlyTrrPheLeuValAlaGlnGlnLeuIleAsnProPheGlyGlu 300
 Db 907 TTTCTTCTTATGTTGGTGTGTTGAAGGTGGCAGAGACGCTCATCAACCCGTTTGGAGAG 966
 Qy 301 AspAspAspAspPheGluThrAsnTrrIleValAspArgAsnLeuGlnValSerLeuLeu 320
 Db 967 GACCATGATGATTTTGGACCAATGATGTGACAGAAATTTGCAAGTCTCCCTGCTG 1026
 Qy 321 AlaValAspGluMetHisGlnAspLeuProArgMetGluProAspMetTyrTrrAsnLys 340
 Db 1027 GCTGTGATGATGATGACACAGGACCTGCTGATGAGAGCCAGCATGTACTGGAAATGAG 1086
 Qy 341 ProGluProGlnProProTyrTrrAlaAlaSerAlaGlnPheArgArgAlaSerPheMet 360
 Db 1087 CTTGAGCACACACCCCTTACAGAGCTGCTTCCGCCAGTTCCGTCAGGCTCTTTATG 1146
 Qy 361 GlySerThrPheAsnIleSerLeuAsnLysGlnGlnMetGlnPheGlnProAsnGlnGlu 380
 Db 1147 GGCCTCCACCTTCAACATCAGCTGACAAAGAGAGATGAGTTTCAGCCCAATCAGAG 1206
 Qy 381 AspGluGlnAspAlaHisAlaGlyTleIleGlyArgPheLeuGlnLysSerHisAsp 400
 Db 1207 GACAAAGAGAGCGCTCACACTGGCATATGGCGGCTTCTTACGGGCTGACGCTCCATGAT 1266
 Qy 401 HisHisProProArgAlaAsnSerArgThrLysLeuLeuTrrProLysArgLysSerLeu 420
 Db 1267 CACCATCTCTCCGGGGGCAAACTCAAGSACCAACTACTGTGGCCCAAGAGGAAATCCCTT 1326
 Qy 421 LeuHisGluGlyLeuProLysAsnHisLysAlaAlaLysGlnAsnValArgGlyGlnGlu 440
 Db 1327 CTCACAGAGGCGCTGCCCAAAACCAACAGGCGCTCAAAAGAAAGCTTAAAGGCGCTGGA 1386
 Qy 441 AspAsnLysAlaTrrPheLysLeuLysAlaValAspAlaPheLysSerGlyProLeuTyrGln 460
 Db 1387 GACAAAGAGGCGCTGAAAGCTTAAAGCTGTAGATGCTTCAAGTCTGCCCACTGTATCAG 1446

Qy 461 ArgProGlyTyrTyrSerAlaProGlnThrProLeuSerProThrProMetPhePhePro 480
 Db 1447 AGGCGAGGCTACTACAGATGGCCCCACAGACGCCCTCTCAGCCCACTCCCATGTTCTTCCC 1506
 Qy 481 LeuGluProSerAlaProSerLysLeuHisSerValThrGlyIleAspTrrLysAspLys 500
 Db 1507 CCAAGACCATCAGGCTGCTCAAGACTTCAACAGTGTCAAGAGCATGACACCAAGACAAA 1566
 Qy 501 SerLeuLysThrValSerSerGlyAlaLysLysSerPheGlnLeuLeuSerGlnSerAsp 520
 Db 1567 AGCTTAAGAAGCTGAGATCTTGGGGCCCAAGAACGTTCTGAATGCTCTCAGAGGAGCAT 1626
 Qy 521 GlyAlaLeuMetGlnHisProGluValSerGlnValArgArgLysThrValGluPheAsn 540
 Db 1627 GGGGCTTGATGAGACCCAGAAATATCTCACGTGAGGAGGAAACCCGTGAGGTTTAAC 1686
 Qy 541 LeuThrAspMetProGluIleProGluAsnHisLeuLysGlnProLeuGlnGlnSerPro 560
 Db 1687 CTGACGATATGCCAGAGATCCCCGAAATCATCTCAAGAACCTTTGGAACTATCAACA 1746
 Qy 561 ThrAsnIleHisThrThrLeuLysAspHisMetAspProTrrTrrAlaLeuGlnAsnArg 580
 Db 1747 ACCAACATACATGCTCACTAAAGATCAAGTGATCTTATTGGGCGCTTGAAGAAACAG 1806
 Qy 581 AspGluAlaHisSer 585
 Db 1807 GATGAAGACATATTC 1821

RESULT 10
 BC041664
 LOCUS
 DEFINITION
 Homo sapiens vitelliform macular dystrophy (Best disease, bestrophin), mRNA (CDNA clone MGC:47884 IMAGE:5194649), complete cds.
 BC041664
 BC041664.1 GI:27371319
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)

REFERENCE
 AUTHORS
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schlier, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marcusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Walek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hilyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grumod, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Small, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL
 MEDLINE
 PUBMED
 12477932
 2 (bases 1 to 2435)
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 Submitted (20-DEC-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK
COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nigr.nih.gov
Ahter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Latic,P., Legaspi,R.,
Mazuro,Q.L., Masello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,
McGowan,J., Pearson,R., Stanciripod,S., Thomas,P.J., Touchman,J.W.,
Taurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
Series: IRAK Plate: 82 Row: d Column: 4.
Location/Qualifiers

FEATURES
source

1. .2435
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:47884 IMAGE:5194649"
/issue_type="Brain, adult, 6 pooled whole brains"
/clone_id="NIH_MGC_114"
/lab_host="DH10B"
/note="vector: pCMV-SPORT6"

gene
1. .2435
/gene="VMD2"
/note="synonyms: BMD, BEST"
/db_xref="locusID:7439"
/db_xref="MIM:153700"
87.11901
/gene="VMD2"
/codon_start=1
/product="VMD2 protein"
/protein_id="AAH41664.1"
/db_xref="GI:27371320"
/db_xref="locusID:7439"
/db_xref="MIM:153700"

CDS
/translation="MPEKLTLYCDSYQIPIPISEVLGRYYVTLVTRWMNOYENLPMD
RLMSLVGFVEGDEQGLRLRLIRYANLGNVLIRSVTAAYKRPSPAOHLVQAGF
MTPAEHQLEKLSLPHNMFVPMVPMFANLSKAWLGGIRIDPILQSLNEMTLQAG
CGHLYADWTSILVYQVYVAYVSPFLICVGRPLNPKAKYRGHELDLVVPPF
LQFFVYGMILKVAQOLINPGEEDDDDEFTNKTVDRLNQLAVDENHODLPMEPDM
YWNKPEQPPYTAASQFRVRSFMSGFENSLNKEEMFQPHQDESDALAGIGRL
GLQSHDHPPLANSRTKLMPKRESLPHLEGIPKHKAKAKOVROQEDNKMKLVAVAK
FKSAPLYORPGYSAQPTPLSPMPFPLPSAPSKLSVTGIDTKDLSKLVTSQAK
KSPFLIESDGLAMHPESQVRKRTVETDMPKLPENHLKRLPLOSPPNITHTLK
DHMPRYALENRSYLHNOGHCIACLRPASLALPFLNHLGFLHNCQSTLDLRPAL
AMGILYATFTGILKCSGPFILSPMHPEDFLGPEGR"

misc_feature
90. .1022
/gene="VMD2"
/note="DUF289; Region: Putative membrane protein. This
family called family 8 in, may be a transmembrane protein
The specific function of this protein is unknown"
/db_xref="CCD:pfam01062"

ORIGIN

Alignment Scores:

Pred. No.: 2,94e-240 Length: 2435
Score: 2817.00 Matches: 527
Percent Similarity: 98.87% Conservative: 0
Best Local Similarity: 98.87% Mismatches: 6
Query Match: 90.29% Indels: 0
DB: 9 Gaps: 0

US-09-622-964A-3 (1-585) X BC041664 (1-2435)

Qy	52	LeuAlaLeuPhrGluGluGlnGlnLeuMetPheGluLeuThrLeuTyCysAspSer	71
Db	60	CTGGCCCTCAGGAAAGAACACAGCTATGTTTAAAGAACTGATGATGGACAGC	119
Qy	72	TyrIleGlnLeuIlePhePhePheValLeuGlyPheTyValThrLeuValThr	91
Db	120	TACATCAGCTCATCCCATTTCTCTGCTGGCTTTCATCGATGAGCTGCTGAC	179
Qy	92	ArgTrpTrpAsnGlnTyGluAsnLeuProTrpAspArgLeuMetSerLeuValSer	111
Db	180	CGCTGCTGCAAGGCAAGACAGACAGAGCCGCTGCTGGCGGACAGCTCATCGGTC	239
Qy	112	GlyPheValGluGlyValAspGluGlnSerArgLeuLeuArgTrpLeuIleArgTyr	131
Db	240	GCGTTCCTCAAGGCAAGACAGACAGAGCCGCTGCTGGCGGACAGCTCATCGGTC	299
Qy	132	AlaAsnLeuGlyAsnValLeuIleLeuArgSerValSerThrAlaValTyrLeuArgPhe	151
Db	300	GCCAACTGGGCAACGCTCATCTCTGCGACGCTCAGACCGGACGCTACAGCGCTTC	359
Qy	152	ProSerAlaGlnIleLeuValGlnAlaGlyPheMetThrProAlaGlnIleGlyGlnLeu	171
Db	360	CCCAAGCCGCGACAGCTGTCGACAGAGCTTTATGACTCCGGCAGAACACAGACATTG	419
Qy	172	GluLeuSerLeuProIleAsnMetPheTrpValProTrpValTyrPheAlaAsnLeu	191
Db	420	GAGAACTGAGCCCTACACACACACATGTTCTGGGTGCTGGGTGGTGGTGGTGGTGG	479
Qy	192	SerMetLeuAlaTrpLeuGlyGlyArgIleArgAspProIleLeuLeuGlnSerLeuLeu	211
Db	480	TCAATGAGCGGTGCTTGGAGGTGCAATCCGGACCTTCTGCTCCAGAGCTGCTG	539
Qy	212	AsnGluMetAsnThrLeuArgThrGlnCysGlyIleLeuTyAlaTyrAspTrpIleSer	231
Db	540	AACGAGATGAACACCTTGCGTACATGATGACACCTGATGCTGCTGCTGCTGCTG	599
Qy	232	IleProLeuValTyrThrGlnValValThrValAlaValTyrSerPhePheLeuThrCys	251
Db	600	ATCCCACTGGTGTATACAGAGGTGTACGTGGCGGTGACACTTCTTCTTCTTCTTCT	659
Qy	252	LeuValGlyArgGlnPheLeuAsnProAlaValAlaTyrProGlyIleGluLeuAspLeu	271
Db	660	CTAGTGGCGCGGCGGCTTCTGACCCAGCCAGCCGCTTACCTGGCCATGAGCTGAC	719
Qy	272	ValValProValPheThrPheLeuGlnPhePheThrValGlyTrpLeuValAla	291
Db	720	GTTGTGCCGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	779
Qy	292	GluGlnLeuIleAsnProPheGlyGluAspAspAspPheGluThrAsnTrpIleVal	311
Db	780	GAGAGCTCATCAACCCCTTGGAGAGATGATATATTTTGAGACCACTGATTTGTC	839
Qy	312	AspArgAsnLeuGlnValSerLeuLeuAlaValAspGluMetHisGlnAspLeuProArg	331
Db	840	GACAGAAATTTGACGGGTGCTCTTGTGCTGTGATGATGACCCAGAGCTGCTGCG	899
Qy	332	MetGluProAspMetTyrTrpAsnIleProGluProGluProProTyrTyrThrAlaSer	351
Db	900	ATGAGGCGGAGCATGTACTGGAATTAACCCGAGCAGACGCCCTTACAGAGCTTCC	959
Qy	352	AlaGlnPheArgArgAlaSerPheMetGlySerThrAsnIleSerLeuAsnIleGlu	371
Db	960	GCCAGTTCGTCGAGCTCTTATGAGCTTCACTTCAACATCACTGAGAACAAAGAG	1019
Qy	372	GluMetGluPheGlnProAsnGlnIleAspGluGluAspGluIleAlaIleGlyIleGly	391
Db	1020	GAGATGAGTTCACGCCCATCAGAGAGAGAGAGATGCTCACGCTGGCATATTGGC	1079
Qy	392	ArgPheLeuGlyLeuGlnSerHisAspHisHisAspProArgAlaAsnSerArgThrLeu	411
Db	1080	CGCTTCTTACGCTGAGCTCCCATGATCACTCTCCAGAGGCAACTCAAGAGACCAA	1139


```

QY 412 LeuLeuTrpProLysArgGluSerLeuLeuHISGluGluLeuProLysAsnHisIleValA 431
Db 1140 CTACGTGGGCCCAAGAGGGAATCCCTTCTCCACAGAGGCGCTGCCCAAAAACCAAGAGCA 1199
QY 432 AAlaYsGlnAsnValAlaGlyGlnGluAspAsnLysAlaTrpLysLeuLysAlaValA 451
Db 1200 GCCAAAGCAAGACGTTAGGGGCGACAGAGACAAACAGCGCTGGAAGCTTAAAGCTGTGGAC 1259
QY 452 AAlaPheLysSerGlyProLeuLysTrpGlnArgProGlyLysTrpSerAlaProGlnTrp 471
Db 1260 GCCTTCAAGCTGCGCCCACTGATCAGAGGCGAGGCTACTACAGTCCGCCACAGAGCGCC 1319
QY 472 LeuSerProThrProMetPhePheProLeuGluProSerAlaProSerLysLeuHisSer 491
Db 1320 CTCAGCCCCCACTCCCATGTTCTTCCCTCGAAGCAATCAGCGCGCTCAAGCTTCACT 1379
QY 492 ValThrGlyLysLeuSerTrpLysAspLysSerLeuLysThrValSerSerGlyLysLys 511
Db 1380 GTCACAGGCAATAGACACCAAGACCAAGAGCTTAAAGACTGTGAGTTCTGGGCGCAAAA 1439
QY 512 SerPheGluLeuLeuSerGluSerAspGlyAlaLeuMetGluHisProGluValSerGln 531
Db 1440 AGTTTGAATTGCTCTCAGAGACCAATGGGCGCTTATGAGAGCCACCAAGATATCTCA 1499
QY 532 ValArgArgLysThrValGluPheAsnLeuThrAspMetProGluLysProGluAsnHis 551
Db 1500 GTAGAGAGGAAACCTGTGAGTTTAACTGACGAGATATCCAGAGATCCCGCAAAATCAC 1559
QY 552 LeuLysGluProLeuGluGlnSerProThrAsnLysHisThrLeuLysAspHisMet 571
Db 1560 CTCAAAGAACTTGTGAGAACATCACCACCAACCAATACACACTACACTCAAGATCACATG 1619
QY 572 AspProLysTrpAlaLeuGluAsnArgAspGluAlaHis 584
Db 1620 GATCCTTATGGGCTTGGAAAAAGTCTGTCTCCAC 1658

```

RESULT 11
LOCUS BC015220
DEFINITION Homo sapiens vitelliform macular dystrophy (Best disease, bestrophin), mRNA (CDNA clone MGC:17761 IMAGE:3877806), complete cds.
ACCESSION BC015220
VERSION BC015220.2 GI:34783515
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Strausberg, R.L., Feingold, B.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heide, F., Dietchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stedelson, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullan, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Wolter, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hui, L.S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whitling, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Schmechel, A., Schein, J.E., Jones, S.J., Skalska, U., Smalins, D.E., human and mouse cDNA sequences
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
MEDLINE 22388257

PUBMED 12477932
REFERENCE 2 (bases 1 to 2441)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (01-OCT-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 1A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT On Sep 16, 2003 this sequence version replaced gi:21955361.
 Contact: MGC help desk
 Email: gcgaps-remail.nih.gov
 Tissue Procurement: DCTD/DTP/Gazdar
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www.hgc.stanford.edu>
 Contact: (Dickson, Mark) mcd@paxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

FEATURES
SOURCE
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: <http://image.llnl.gov>
 Series: IRAC Plate: 14 Row: b Column: 9.
 Location/Qualifiers
 1..2441
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:17761 IMAGE:3877806"
 /tissue_type="Lung, large cell carcinoma"
 /clone_lib="NIH_MGC_68"
 /lab_host="DH10B"
 /notes="Vector: pCMV-SPORT6"
 1..2441
 /gene="VMD2"
 /notes="synonym: BMD, BEST"
 /db_xref="locusID:7439"
 /db_xref="MIM:153700"
 573..2069
 /gene="VMD2"
 /codon_start=1
 /product="VMD2 protein"
 /protein_id="AAH15220.1"
 /db_xref="GI:22800661"
 /db_xref="locusID:7439"
 /db_xref="MIM:153700"
 /translation="MFEKLTLYCDSTYLTOLIPISFVLGFVYLVVTRWNOYENLPMPDRLMSIVSGFVGKDEQRLRLRLRYANIGNVILRSVSTAVYKRPSPAOHLVOAGFMTPEAHKQLEKLSLPHNMFVPMVPMFANLSKMLGRIIDPILLOSILNEMNTPTROCGHYAVDWISIPLYTVQVTVAVYSPFLICVAGROINPAKAYRGEHLDLVVPTFENILNENEMFQNOEDEBDAHAGILGRFLGQSHDHPRANRSRTKLMPRESILHEGILPKHAKQNVNRQEDENKAKWAKVADAFSAQPYQRPYVSAPOPTLSPTPMFPLESAPSKHAKVQVNDIKDLSIKTVSSGAKSEFELSESDGALMHEPVSQVRRRTVBFNLTIDPEIPEENILKEPLEOSPNTHTLKHMDPYWALENDEAHS"
 576..1427
 /gene="VMD2"
 /notes="ndup289; Region: Putative membrane protein. This family called family 8 in, may be a transmembrane protein The specific function of this protein is unknown"
 /db_xref="CDD:pfam01062"

misc_feature
ORIGIN
 Alignment Scores:
 Pred. No.: 2,096-227 Length: 2441
 Score: 2672.50 Matches: 505
 Percent Similarity: 94.57% Conservative: 0
 Best Local Similarity: 94.57% Mismatches: 2
 Query Match: 85.66% Indels: 27
 DB: 9 Gaps: 1

US-09-622-964A-3 (1-585) x BC015220 (1-2441)

OY	52	LeuAlaLeuThrGluGluGlnGlnIleuMetPheGlyIuyluPheThrLeuTYrCYAspSer	71
Db	546	CTGGCCCTCAGGAAACAAACAGCTGATGTTTGAAACATGACCTGTATATGGCAAGC	605
OY	72	TyrIleGlnLeuIleProIleSerPheValLeuGlyPheTYrValThrLeuValAlaThr	91
Db	606	TACATCCAGCTCATCTCCATTTCTTCGTGGCTGGCTTCTATGTAACCTGGTCTGAC	665
OY	92	ArgTrpTPraenGlnTYrGluAsnLeuProTrpProAspArgLeuMetSerIleuValSer	111
Db	666	CGCTGGTGAACCAAGTACAGAACTGGCCGTGGCCCGACCGCTCATGAGCTGGTGTGCG	725
OY	112	GlyPheValGluGlyIuAspGlnIleSerArgLeuLeuArgTYrThrLeuIleArgTYr	131
Db	726	GACCTTCGTGAAAGGCAAGGACAGGACAGCCGAGCTGCGGCGGACGTCATCCGCTAC	785
OY	132	AlaAsnLeuGlyAsnValLeuIleLeuArgSerValSerThrAlaValTYrIlyAspPhe	151
Db	786	GCCAACTGGGGCAACGTGCTCATCTCGGACGCGTCAACACCGGACGTCTACAGAGCTTC	845
OY	152	ProSerAlaGlnIleLeuValGlnAlaGlyPheMetThrProAlaGlnIleIlyGlnLeu	171
Db	846	CCCAAGCCGCACAGCACTGTGTGCACAGACGCTTTAAGTCCGGCAAGAACAAAGCATGG	905
OY	172	GluIlyIleuSerLeuProHisAsnMetPheTrpValProTrpValTrpPheAlaAsnLeu	191
Db	906	GAGAAACTGAGCCTTACCAACAACTGTTCTTGCGTGGCCCTGGGTGTGGTTGCCAACTG	965
OY	192	SerMetIlyValAlaTrpLeuGlyGlyValArgIleArgAspProIleLeuLeuGlnSerLeu	211
Db	966	TCAATGAAGAGCGGTGGCTTGGAGGTGCAGATCCGGGACCCATTCCTGCTCCAGAGCTGCTG	1025
OY	212	AsnGluMetAsnThrLeuArgThrGlnCysGlyIleuTYrAlaTYrAspTrpIleSer	231
Db	1026	AACGAGATGAACACTTGGCTGATCTAGAGTGTGACACCTGTATGCTTACAGACTGATTAAGT	1085
OY	232	IleProLeuValTYrTrhGlnValAlaThrValAlaValTYrSerPhePheLeuThrCys	251
Db	1086	ATCCCACTGGTGTATACACAGGTGTGACTGAGCGGATGACAGCTTCTTCCTGACTTGT	1145
OY	252	LeuValGlyIuArgGlnPheLeuAsnProAlaIlyValAlaTYrProGlyHisGluLeuAspLeu	271
Db	1146	CTAGTGTGGCGGCACTTCTCTGAACCCACCAAGGCTTACCTCGGCAAGAGCTGGACCTC	1205
OY	272	ValValProValPheThrPheLeuGlnPhePhePheTYrValGlyTrpLeuIlyValAla	291
Db	1206	GTTGGCCCGCTTCAAGTCTCTGCAAGTCTTCTTCTATGTTGGCTGACTGAAG-----	1255
OY	292	GluGlnLeuIleAsnProPheGlyIuAspArgAspArgPheGluThrAsnTrpIleValAl	311
Db	1259	-----	1259
OY	312	AspArgAsnLeuGlnValSerIleuLeuAlaValAspGluMetHisGlnAspLeuProArg	331
Db	1260	-----GATGCCCTGTGGCTGTGGATGAGATGACAAAGAACCTGCTGGG	1304
OY	332	MetGluProAspMetTYrTrpAsnIlyuAspGluProGluProGlnProProTYrTrhAlaAlaSer	351
Db	1305	ATGAGAGCGGACATGTACTGGAATAAAGCCGAGCCACAGCCCCCTTACACAGCTGCTTCC	1364
OY	352	AlaGlnPheAlaGlyAlaSerPheMetGlySerThrPheAsnIleSerLeuAsnIlyGln	371
Db	1365	GCCCAAGTTCGTCGAGCCTCTCTTAATGGAGCTCCACCTTCAACATCAGGCTGAACAAAG	1424
OY	372	GluMetGluPheGlnProAsnGlnIuAspGlnIuAspAlaAlaAlaGlyIleIleGly	391
Db	1425	GAGATGAGATTCCAGCCCAATGAGGAGACAGAGAGATGCTCAAGCTGGCATCATTTGGC	1484
OY	392	ArgPheLeuGlyLeuGlnSerHisAsnHisIlyuProAlaGlyAlaAsnSerArgThrIlyu	411
Db	1485	CGCTTCCATGAGCTGCAAGTCCATATATCAACATCTCTCCAGGAGCAATCTCAAGAGCCAA	1544

OY		412	LeuEUTPRPOLYVAAGLUSERLEUEHISGLUYLEUPROLYVARNILYLVALA	431
Db		1545	CTACTGTGGCCCAAGAGAATTCCTTCTCACAGGCGCTGCCAAAACSCAAAGCSA	1604
OY		432	AlAlyeGIAnPaValArGVLGNIGLYNRYANLYYVAIATPLYLVEULYALAVAIAPR	451
Db		1605	GCCAAACAAGACGTTTAGGGCCAGBMAACAACAAGCCTTGAAAGCTTAAGCTGTGCAC	1666
OY		452	AlAheLYSeSGlYPProLeuTYrGLINARPrOGLIYYTYrYSERAlAProGLINrPro	471
Db		1665	GCGTTCAAGTCTGCCCACTGTATACAGAGGCSAAGCTACTACAGTCCGCCAACACACC	1724
OY		472	LeuSerPROTHrPROMerPherPharProLeugLynrOSerAlAProSerLYeUHISeS	491
Db		1725	CTCAGCCCCACCTCCCATGTTCTTCCCCTTAAMACCATACAGGCCGTCAAAGCTCACAGT	1784
OY		492	VAlThrgLIleApTrnLYaSPrLYaSPrLYeSerLeuLYSThrVALeSerGdLYAlAYelYe	511
Db		1785	GTCACAGGCAATAGACACCAAAAGACMAAAGCTTAAAGACTGTGAGTCTGGGGCCAAAGAA	1844
OY		512	SerPhegLuleuLeuSerGluSerAepGLYAlAleuWetGUHIsPROGLINAlSeSLn	531
Db		1845	AGTTTGAATTGCTCTCAGAGAGCGATGGGGCTTGATGGAGCACCCAGAAAGTATCTCAA	1904
OY		532	VAlArgaGLySThrVALGlUpHpaAnLYeThrArMetProGUILleProGLINAHnHs	551
Db		1905	GTCGAGAGAAACCTGTGAGACTTTAACTGACGATATGCCAGATATCCGAAAGATCAC	1964
OY		552	LeuLYeGluPProLeuGluGlnSerProThraNIleHISThrTrnLeuLYaSPRIseWc	571
Db		1965	CTCAAGAACCTTTGGAAACATCAACCAACCAACATCACTACACTCAAAAGATCACTAG	2024
OY		572	AapPROTYrTPDAlAleuGLIAenNArGaSPGLIALIHISer	585
Db		2025	GATCCTTAATGGCGCTTGAAAAACAGGAGATGAACACACTTCC	2066
RESULT 12				
BC079048				
LOCUS	BC079048	2035 bp	mRNA	linear ROD 03-AUG-2004
DEFINITION	Rattus norvegicus cDNA clone MGC:93985 IMAGE:J115295, complete cds.			
ACCESSION	BC079048			
VERSION	BC079048.1	GI:50927692		
KEYWORDS	MGC.			
SOURCE	Rattus norvegicus (Norway rat)			
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
REFERENCE	Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Bluetow, K.H., Schaefer, C.F., Bhac, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Halel, F., Ditschenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Sieblton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Schneez, T.E., Brownstein, M.J., Udell, T.B., Tohyuki, S., Caminci, P., Prange, C., Kaha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosok, S.A., McEwan, P.J., McMernan, K.C., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulik, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahney, J., Helton, E., Ketteman, M., Medan, A., Young, A.C., Rodighiero, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bonfard, G.G., Blackesley, R.W., Touchman, J.T., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimm, J., Schmutz, U., Myers, R.M., Butcherfield, Y.S., Krzywinski, J.M.I., Skalska, U., Smalins, D.E., Scherich, A., Schein, J.E., Jones, S.U. and Marx, M.A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences			
JOURNAL	Proc. Natl. Acad. Sci. U.S.A.	99 (26),	16899-16903	(2002)
PUBMED	12477932			
REFERENCE	2 (bases 1 to 2035)			

OY	420	leuLeuH16GIUGLyleuProlysernH1slybA1a1alyGlnAserN1AArglyGln	433
Db	1413	TTCTCC---GGAGACCACTCTAAGAATACAGACGAAAGCCAAATA-----	1457
OY	440	GlnAerAenlybA1a1rlybLeuylbA1aValAa1aPhelySerGlyProleuTyr	459
Db	1488	-----AATGATGTCTGGAAACCAAG-----GACCTCTTCAAGACTATACCAAGTTT	1505
OY	460	GlnArgProGlyTyrTyrSerAlaProGlnThrProLeuSerProThrProMetPhePhe	479
Db	1506	AAGAGCCCAAGTTACCGCAGGCCCCACAGCCACTTGGCAGCCACTTCCCATGGTCTTC	1565
OY	480	ProLeuGlnProSerAlaProSerLysLeuH1SerValThrGlyAlaSprThylAsp	499
Db	1566	CCTCCTCAGCAGTCCAGCACCCCTCAGTGGCTTACTACGACACAGGT-----	1610
OY	500	LysSerLeuLysThrValSerSerGlyAlaLysLysSerPheGlnLeuSerGluSer	519
Db	1610	-----	16110
OY	520	AerGlyAlaLeuMetGlnH1aProGluValSerGlnValArgArgLysThrValGlnPhe	539
Db	1611	GTTGGGCCCTTCCGCACACTACCAAGAGCCCTGTCTACATGAAGAAACAGTGGAGTTT	1670
OY	540	AenLeuthrAerPmetProGlnL1aProGlnbH1sLeuLysGlnProLeuGlnGlnSer	559
Db	1671	AACTTGG---AACATTCGACAGAGTGGCCCTACGGAACATGAACAAACCCATTTGGACAG---	1724
OY	560	ProThrAenH1eH1ThrThrLeuLysAerH1sMetAerProTyrTyrAlaLeuGlnbN	579
Db	1725	-----ATGCACACGATCTCTGAAGAGCCTGGAGAACTCATCTTCTAGTGATAA	1775
OY	580	ArgAer 581	
Db	1776	GCTGAC 1781	
RESULT 13			
LOCUS	BD136744	1916 bp	DNA
DEFINITION	Beet's macular dystrophy gene.		
ACCESSION	BD136744		
VERSION	BD136744.1	GI:23231689	
KEYWORDS	JP 2002504559-A/26.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 1916)		
AUTHORS	Petrukhin, K., Caskey, T. C., Metzker, M. and Wadelius, C.		
TITLE	Beet's macular dystrophy gene		
JOURNAL	Patent: JP 2002504559-A 26 12-FEB-2002; MERCK & CO INC, CLAES WADELIIUS		
COMMENT	OS Mus musculus (mouse) PN JP 2002504559-A/26 PD 12-FEB-2002 PF 22-FEB-1999 JP 2000533447 PR 25-FEB-1998 US 60/07541.18-DEC-1998 US 60/112926 PI KONSTANTIN PETRUKHIN, THOMAS C CASKEY, MICHAEL METZKER, CLAES PI WADELIIUS PC C07K16/18, C07K14/47, C12N5/10, C12N15/09, C12P19/34, C12Q1/68// PC C12P21/08, PC C12N5/00, C12N15/00 CC Beet's macular dystrophy gene FH Location/Qualifiers FT source 1..1916 Location/Qualifiers 1..1916 /organism='Mus musculus (mouse)'. 1..1916 Location/Qualifiers 1..1916 /organism='Mus musculus' /mol_type='genomic DNA' /db_xref='taxon:10090'		
FEATURES	source		
ORIGIN			

Alignment Scores:			
Pred. No.:	9,21e-159	Length:	12916
Score:	1899.00	Matches:	378
Percent Similarity:	74.32%	Conservative:	59
Best Local Similarity:	64.29%	Mismatches:	101
Query Match:	60.87%	Indels:	50
DB:	6	Gaps:	10
US-09-622-964A-3 (1-585) x BD136744 (1-1916)			
QY	1 MetThrIleThrTYrThsSerGlnValAlaAsnAlaArgLeuGlySerPheSerArgLeu	20	
DB	11 ATGACTATTCACCTTACACACAAAGTGGCAATGCCGCTCGCTTCGTCTGCTCC	70	
QY	21 LeuLeuCySTrPArgIySerIleTyrlYsLeuLeuTyrcIyGluPheLeuIlePheLeu	40	
DB	71 CTCCTGTGTGGCGAGGACGATCTTCAAGCGTGTGTATGAGAAATTCCTTGTCTTCATA	130	
QY	41 LeuCySTrTYrIleIleArgPheIleTyraGluAlaLeuThrGluGlnGlnLeu	60	
DB	131 TTCCTCTACTATTCATCCGCTGAGACTCTACAGAAATGGTCTCTCGAGATATCGACAGCTG	190	
QY	61 MetPheGluVylveLeuThrLeuTyrcYsAspSerTYrlaGluLeuIleProlIeserPhe	80	
DB	191 TTGTTTGAGAACTGGCTCTGTACTCGACGACTACATTCAGCTCAATCCTTAATCTTTC	250	
QY	81 ValLeuGlyPheTyValThrLeuValValIThrArgTrpAsnGlnTyrcIyAsnLeu	100	
DB	251 GTTCTGGGTTCTTATGTTTACATTTGGTGGTGAAGCCGTGGTGGAGCCAGTACGAACCTTG	310	
QY	101 ProTrpProAspArgLeuMetSerLeuValSerGlyPheValGluGlyLysAspGluGln	120	
DB	311 CGGTGGCCGACCGGCTCATGTATCCAGAGTGTACCTGTGTGAGGGCAAGATGAGGAA	370	
QY	121 SerArgLeuLeuArgArgThrLeuIleArgTYrAlaAsnLeuGlyAsnValLeuIleLeu	140	
DB	371 GGCGCTTTGGTCGGGGGCGACGCTCATCCCTACGCCATCTGGGCAAGTGCATCTCTG	430	
QY	141 ArgSerValSerThrAlaValTYrLysArgPheProSerAlaGlnHisLeuValGlnAla	160	
DB	431 CGCAGCATCAGACCTCGGTCTTACAGGCCCTTCCACCTTTCACCACTCGTGTCTTACGA	490	
QY	161 GlyPheMetThrProAlaGlnHisLysGlnLeuGlyLysLeuSerLeuProHisAspMet	180	
DB	491 GGTTTTATGACCAACCATGGGAACTAAGACAGTTGCAGAACTGGGCTTACACACAACACA	550	
QY	181 PheTrpValProTrpValTrpPheAlaAsnLeuSerMetClyValaTrpLeuGlyGlyArg	200	
DB	551 TTCTGGGTGCGCTGGGTGTGGTTTGGCACTTGTCAATGAGGCTTATGTGAGAGTGA	610	
QY	201 IleArgAspProIleLeuLeuGlnInserLeuLeuAsnGluMetAsnThrLeuAspThrGln	220	
DB	611 ATCCGGGAGAACCGTCTCGTCTCCAGAGCCTGATGAATGAGGTGTGTACTTTGGCACTCAG	670	
QY	221 CysGlyHisLeuTYrAlaIleTYrAspTrpIleSerIleProLeuValTYrThrGlnValVal	240	
DB	671 TGTGACACAGCTGTATGCTTACAGACCTGGAATAGTATCCATTGGTGTATCACAACAGTGGTG	730	
QY	241 ThrValAlaValTYrSerPhePheLeuThrCYsLeuValGlyArgGlnPheLeuAsnPro	260	
DB	731 ACAGTGGCAGTATTAACAGCTTTTCTTGGATGCTTATCGGGAGGACAGTTTTCGAACCCA	790	
QY	261 AlaLysValaTYrProGlyHisGluMetAspLeuValValProValPheThrProLeuGln	280	
DB	791 AACCAAGACTACCCAGGCGCATAGATGATCTGGTTGTCCCTGTCTTCAAAATCTTCGAA	850	
QY	281 PhePhePheTYrValGlyTrpLeuLysValAlaGluGlnLeuIleAsnProPheGlyGlu	300	
DB	851 TTCTTATTTCTACATGGGCTGGCTGAAGGTGGCAACACTATCAACCCCTTCGGGGAG	910	
QY	301 AspAspAspAspPheGluThrAsnTrpIleValAspArgAsnLeuGlnValSerLeuLeu	320	
DB	911 GACGATGTGATTTTGAAGCTAACTAGATCATTTGACAGAACTCGAGGTGTCCCTGTG	970	

```

Qy 321 AlavAlaSerGluMetHisGlnAspLeuProArgMetGluProAspMetTyrTrpAsnLys 340
Db 971 TCCTGGATGGGATGACACCAAGAGTCTCTCCAGACGTCAGTCACTGTAACGAG 1030
Qy 341 ProGluProGlnProProTyrTrpAlaAlaSerAlaGlnPheArgAlaSerPheMet 360
Db 1031 GCACGGCTCAGCGCCCTACAGCGCTTCTCCAGCGCTCGCCGACCTTCTTCTCAG 1090
Qy 361 GlysSerPheAsnLysSerLeuAsnLysGluGluMetGluPheGlnProAsnGlnLys 380
Db 1091 GGCTCCACCTTCACATCAGCTTAAGAGAAAGACTTAAGAGCTTTGGTCAAAAGAGAG 1150
Qy 381 -----AspGluAspAlaHisAlaGlyLeuLeuGlyLeuPheLeuGlyLeu 396
Db 1151 GCTGACACGATAGAAAGAGAGTGTATAGACAGACCTAGAGCTGCTTCTTAGGACTG 1210
Qy 397 GlnSerHisAspHisSerProProArgAlaAsnSerArgThrLysLeuLeuTrpProLys 416
Db 1211 CAACCCMAAACTACATCTTCTCCGMAAAAGACTTAAGACCAACTATTGTGTTCTAG 1270
Qy 417 ArgGluSerLeuLeuHisGluGlyLeuProLysAsnHisLysAlaAlaLysGlnSerVal 436
Db 1271 AACCCCTCTC-----GAGGCCAGTGTAAAGAT-----GCCAACCAAG 1309
Qy 437 ArgGlyGlnGluAspAsnLysAlaTrpLysLysAlaAlaValAspAlaPheLysSerGly 456
Db 1310 AAAAACAACAGAAAGAT-----GTCTGGAATTTAAAGCGTCTGACTTCTGAAATGTT 1363
Qy 457 ProLeuTrpGlnArgProGlyTyrTyrSerAlaProGlnThrProLeuSerProThrPro 476
Db 1364 CCAAGTTTAAAGAGAGGCTCCCTCTGCTGCGCCACAGCAGCAGCAGCAGCAGCAGCAGC 1414
Qy 477 MetPhePheProLeuGluProSerAlaProSerLysLeuHisSerValThrGlyLysAsp 496
Db 1415 -----CACCTTACTGAGCAGTCAAGACCTCTCACT-----TCAGAACACAGGT----- 1456
Qy 497 ThrLysAspLysSerLeuLysThrValSerSerGlyAlaLysLysSerPheGlnLeuLeu 516
Db 1456 ----- 1456
Qy 517 SerGluSerAspGlyAlaLeuMetGluHisProGluValSerGlnValArgArgLysThr 536
Db 1457 -----GATGGCCCTCCACAGATTTACCAAGAAATCTGTCACTCAAAAAAAACT 1507
Qy 537 ValGluPheAsnLeuThrAspMetProGluLeuLeuProGluAsnHisLysLysGlu---Pro 555
Db 1508 GTGAGGTTTAACTTG---ACATTCACAGAGAGCCCAAGAACATCTTCAACAGCGCGGT 1564
Qy 556 LeuGluGlnSerProThrAsnLysHisThrThrLeuLysAspHisMetAspProTyrTrp 575
Db 1565 TTGACCCAGATGTCAACCAATATACAGGCTTAAATGAAGAGCATGACAGTCTAT--- 1621
Qy 576 AlaLeuGluAsnArgAspGluAla 583
Db 1622 -----CCCTACAGGATGAAGCT 1639
RESULT 14
AY450427 1904 bp mRNA linear ROD 24-JUN-2004
LOCUS Mus musculus vitelliform macular dystrophy 2 (Vmd2) mRNA, complete
cDNA.
ACCESSION AY450427
VERSION AY450427.1 GI:41612986
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1904)
Stohr, H., Marguard, A., Nanda, I., Schmidt, M. and Weber, B.H.
Three novel human VMD2-like genes are members of the evolutionary
highly conserved RFP-TM family

```

```

JOURNAL Eur. J. Hum. Genet. 10 (4), 281-284 (2002)
MEDLINE 22027749
PUBMED 12032738
REFERENCE 2 (bases 1 to 1904)
AUTHORS Kraemer, F., Stohr, H. and Weber, B.H.F.
TITLE Cloning and characterization of the murine Vmd2 RFP-TM gene family
JOURNAL Cytogenet. Genome Res. 105 (1), 107-114 (2004)
REFERENCE 3 (bases 1 to 1904)
AUTHORS Weber, B.H.F. and Kraemer, F.
TITLE Direct Submission
JOURNAL Submitted (27-OCT-2003) Human Genetics, University of Wuerzburg, Am
Hubland, Biocentre, Wuerzburg 97074, Germany
FEATURES
source
1..1904
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/chromosome="19"
/tissue_type="testis"
1..1904
/gene="Vmd2"
37..1692
/gene="Vmd2"
/codon_start=1
/product="vitelliform macular dystrophy 2"
/protein_id="AA09922.1"
/db_xref="GI:41612989"
/translacion="MTITVTKVANAARIGSFSSLLICWRGSIYKLLGEPVFTLY
SIRGLRYMTLSDDOQLPEKALYCDYVGLPIPIFYAGFYTTLVSRMGSYENLPW
PDRMTIOVSPFVGEDEGRRLRLIRAILIGVLIIRSTISYKRPPTLHLVLA
GPMTREHQLQKLGPHNTEVWVWVWVFNLSMKAYLGGRIKDTVLQSLNVECTLR
TQGGQLYADWISIPLVYTVVAVVAFELACLIGQFIPNPKDYPGHMDLVVPL
TLIQFLFYMGLKVAEOLINPGEHDDDEFTWILIDRMLOVLSLSDGMHONLPMER
DYMWEAEPQPYTAASARSRRHSFMGSTPINLSKEDLELMSKEADTDKRGYS
TIGCFGLQPKYHLPLKDLTKLCSKNPLLEGQKANOAKNOVDWKRFGDPLKC
VPRFRGSHGCPQAPSSHPTEQSA PSSSDPGDSTPYQGI CHMKKKTVEPNLIPB
SPTEHLQQRRLDQNSTNIGALMKHAESEYPRDEAGTKEVLYE"
ORIGIN
Alignment Scores:
Pred. No.: 1,69e-158 Length: 1904
Score: 1896.00 Matches: 377
Percent Similarity: 74.32% Conservative: 60
Best Local Similarity: 64.12% Mismatches: 101
Query Match: 60.77% Indels: 50
DB: 10 Gaps: 10
US-09-622-964A-3 (1-585) x AY450427 (1-1904)
Qy 1 MetThrIleThrTyrTrpSerGluValAlaAsnAlaArgLeuGlySerPheSerArgLeu 20
Db 37 ATGACTATACACTTACACAAACAAGATGACCAATCCCGCTCGGTGTTCTGCTCCCTC 96
Qy 21 LeuLeuGlySerTrpArgLysSerLysLysLeuLeuTyrGlyGluPheLeuLeu 40
Db 97 CTCCTGCTGCGCAGGAGCAGCATCAACAGTGTGTATGAGAAATCTTGTCTTATATA 156
Qy 41 LeuGlyTyrTrpLysLysArgPheLysLysLeuAlaLeuThrGluGlnGlnLeu 60
Db 157 TTCCTCTACTATTCATCCATCCGAGCTCTACAGATGTTCTCTCCAGTCAAGCAGCTG 216
Qy 61 MetPheGluLysLeuThrLysLysCysAspSerTyrTrpGlnLeuLeuPheLeuSerPhe 80
Db 217 TTGTTTGAAGAGCGGCTCTGTACTGCGACAGCTACATCCAGCTCATCCCTATATCTTC 276
Qy 81 ValLeuGlyPheTyrValThrLeuValAlaThrArgTrpPheAsnGlnTyrGluAsnLeu 100
Db 277 GTTCTGGGTTTCTATATGTTACATGTTGTGTGAGCGCGCTGGTGGAGCCAGTACGAACTTG 336
Qy 101 ProTrpProAspArgLeuMetSerLeuValSerGlyPheValGluGlyLysAspGluGln 120
Db 337 CCGTGGCCCAACCGCTCATGATCCAGGTGTCTTACGCTGTGTGAGGCAAGATGAGAA 396

```

```

Qy 121 SerArgLeuLeuArgArgThrLeuIlleArgTyrAlaIleuLeuGlyAsnValLeuIlleu 140
Db 397 GCGCGTTGTCGCGCGCGACCTCATCCGCTACCGCATCCCGCGCAAGTCTCATCTCG 456
Qy 141 ArgSerValSerThrAlaValTyrIleValArgPheProSerAlaGlnHisLeuValGlnAla 160
Db 457 CGCAGCATCAGACCTCGGTCTACAGAGCGTTTCCCATCTTCCACCTGCTGCTGACAG 516
Qy 161 GlyPheMetThrProAlaGlnHisIleValGlnLeuGluValLeuSerLeuProHisAsnMet 180
Db 517 GGTITTTAGACCTATGGGGAACATAGCATGGAAGTGGGCTTACCAACACACACA 576
Qy 181 PheTyrValProTyrValTyrPheAlaAsnLeuSerMetLeuValTyrIleValGlyArg 200
Db 577 TTTCTGGGTGCGCTGGGTGTGTTTCCCACTTGTCAAGAAAGCGCTTCTTGAGGTGCA 636
Qy 201 IleArgAspProIleLeuLeuGlnSerLeuLeuAsnGluMetAsnThrIleuArgThrGln 220
Db 637 ATCCGGGACACCGTCTGCTCCAGAGCGCTGATGATGATGATGATGATGATGATGATGAT 696
Qy 221 CysGlyHisLeuTyrAlaTyrAspTyrIleSerIleProLeuValTyrThrGlnValVal 240
Db 697 TGTGACAGCTGTATGCTTACGATGATGATGATGATGATGATGATGATGATGATGATGAT 756
Qy 241 ThrValAlaValTyrSerPhePheLeuThrCysLeuValGlyArgGlnPheLeuAsnPro 260
Db 757 ACAGTGGCAGATACAGCTTTTCTTCCATGCTTGTATCGGGAAGCAAGTTCCTAACC 816
Qy 261 AlaIleValTyrProGlnHisGlnLeuAspLeuValValProValPheThrPheLeuGln 280
Db 817 AACAAAGCATCCACAGCGCATGAGATGATGATGATGATGATGATGATGATGATGATGAT 876
Qy 281 PhePhePheTyrValGlyTyrPheLeuValAlaGlnLeuIleAsnProPheGlyGlu 300
Db 877 TTTCTTATCTTCAATGCGGTGCTGATGAGGTGCAAGACGCTCATCAACCCCTTCG 936
Qy 301 AspAspAspAspPheGluThrAsnTyrIleValAspAspAsnLeuGlnValSerLeu 320
Db 937 GACATGATGATTTTGAAGACTTACATGATCATGATCAAGAAACCTGCAAGGTCTCC 996
Qy 321 AlaValAspGluMetCHisGlnAspLeuProArgMetGluProAspMetTyrTyrAsnLys 340
Db 997 TCCGTGATGGAGTCAACCAACTGCTGCTCCCATGGAAGCTGACATGATGATGATGAT 1056
Qy 341 ProGluProGlnProProTyrThrAlaAlaSerAlaGlnPheArgArgAlaSerPheMet 360
Db 1057 GCAAGCGCTCAGCGCGCTTACAGAGCTGCTTCCAGAGTCTCGCGCATTCCTCATG 1116
Qy 361 GlySerThrPheAsnIleSerLeuAsnHisGlnGluMetGluPheGlnProAsnGlnGlu 380
Db 1117 GGTCTCACTTCAACATCAGCTTAAGAAAGAAACCTTAAGGCTTGTGCTCAAAAGGAG 1176
Qy 381 -----AerGluGluAspAlaHisIleAlaGlyIleIleGlyArgPheLeuGlyLeu 396
Db 1177 GCTGACACGATAGAAAGAGAGTGCCTATAGCAGACCAAGCTGCTTCTTAGAGCTG 1236
Qy 397 GlnSerHisAspHisHisProProArgAlaAsnSerArgThrIleValLeuTyrProLys 416
Db 1237 CAACCCAAAACCTACCTCTTCCCTTGAAGACTTAAGACCAACATGATGTTGTTCTTAAG 1296
Qy 417 ArgGluSerLeuLeuHisGlnGluValLeuProLysAsnHisIleValAlaIleValGlnAsnVal 436
Db 1297 AACCCCTCTCTC-----GAAGGCCAGCTGAAGGAT-----GCCAACCAAG 1335
Qy 437 ArgGlyGlnGluAspAsnLysValAlaTyrPheLeuValValAlaAspAlaPheLysSerGly 456
Db 1336 AAAAACCAAGAAAGT-----GTCTGAAATTTAAAGGTCTGCACTTCTGAATGTGCTT 1389
Qy 457 ProLeuTyrGlnArgProGlyTyrTyrSerAlaProGlnThrProLeuSerProThrPro 476
Db 1390 CCAAGGTTTAAAGAGAGAGGCTCCCATTTGTGCGCCACAGGAGCAGCAGCAGCAGCAGCAGC 1440

```

```

Qy 477 MetPheProLeuGluProSerAlaProSerLysLeuHisSerValThrGlyIleAsp 496
Db 1441 -----CACCTTACTAGACATGACAGACCCCTCCAGT-----TCAGACACAGGT----- 1482
Qy 497 ThrLysAspLysSerLeuLysThrValSerSerGlyAlaLysLysSerPheGluLeu 516
Db 1482 ----- 1482
Qy 517 SerGluSerAspGlyAlaLeuMetGluHisProGluValSerGlnValArgArgLysThr 536
Db 1483 -----GATGGGCGCTTCCACAGATTACCAAGAAATCTGTGCATGATAAAGAAACT 1533
Qy 537 ValGluPheAsnLeuThrAspMetProGluIleProGluAsnHisLeuLysGlu---Pro 555
Db 1534 GTGAGATTAACTTG---AACATTCAGAGAGCGCCCAAGAAACATCTTCAACAGCGCGCT 1590
Qy 556 LeuGluLysSerProThrAsnIleHisThrThrLeuLysAspHisMetAspProTyrTyr 575
Db 1591 TTGGACCAAGATGTCAACCAATATATACAGGCTCTATGTAAGAGCATGACAGTCTCTAT--- 1647
Qy 576 AlaLeuGluAsnArgAspGlyAla 583
Db 1648 -----CCTTACAGGAGATGAAGCT 1665

RESULT 15
AY064707 1289 bp mRNA linear MAM 19-AUG-2002
LOCUS AY064707
DEFINITION Sus scrofa beestrophin mRNA, partial cds.
ACCESSION AY064707
VERSION AY064707.1 GI:18476495
KEYWORDS
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 1289)
Marmorstein, L.Y., McLaughlin, P.J., Stanton, J.B., Yan, L., Crabb, J.W.
and Marmorstein, A.D.
Beestrophin interacts physically and functionally with Protein
Phosphatase 2A
Phosphatase 2A
J. Biol. Chem. 277 (34), 30591-30597 (2002)
12058047
2 (bases 1 to 1289)
Marmorstein, L.Y., McLaughlin, P.J., Stanton, B., Yan, L., Crabb, J.W.
and Marmorstein, A.D.
Direct Submision
Submitted (30-NOV-2001) Cole Eye Institute, 331, Cleveland Clinic
Foundation, 9500 Euclid Avenue, Cleveland, OH 44195, USA
FEATURES
source
1..1289
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
<1..1289
/codon_start=3
/product="beestrophin"
/protein_id="AAL40882.1"
/db_xref="GI:18476496"
/translation="PQHLYVAGFMTPESEHKLKLSLPHNSFMKPMWVPMNLSTKANI
GGRILRPVLLSLDEKMTLTLCGSLYADMSVPLVYQVTVAVSFLACTLVR
OFLNPKAPVAGHMDLVLPFLQFFPYGWLKVALQLINPFGEDDDPEMTNIVR
SLQVRSAYNDEMNDLRPMKRDMMYNDREHPRYTASASQSRSPFGSGTNPSTIGRE
DMERQVEEBEENATGILGNLGLQSSDHQRPRTNSKTLMPKKEGFHNGHNPGLNG
ARLDDSSQDEBSKMRREGFSAALCSRPQVNSAPQVPLGHTPMVFPPEBSAPVGLRVR
SGIDEAKQSLQPATPISIKKSFELLPESAEASAPROGSHVNRKTVIEFNLDSEAP
EHKLEPMLPEPMGIIHALIKDRDPYALLENRDBANS"
ORIGIN
Alignment Scores: 2.2e-141 Length: 1289
Pred. No.: 1701.00 Matches: 324
Score: 81.52% Conservative: 29
Percent Similarity: 74.83% Mismatches: 72
Best Local Similarity:

```

Query Match: 54.52% Indels: 8
 DB: 4 Gaps: 4
 US-09-622-964A-3 (1-585) x AY064707 (1-1289)

```

QY      155 GlnHisLeuValGlnAlaGlyPheMetThrProAlaGlnHisLeuGlnLeuGlnLeu 174
DB      6 CAAACACCTGGTGAAGAAGAGGCTTATAGACCCCTCGGAAACAAGACCTTAGAAAAACCTG 65
QY      175 SerLeuProHisAsnMetPheTrpValProTrpValTrpPheAlaAsnLeuSerMetLys 194
DB      66 AGCTTGCAACAACAACCTATCTGAGATGCTGGGTGGTGGTTTCCAAACCTGTCAACGAAG 125
QY      195 AlaTrpLeuGlnGlyArgIleArgAspProIleLeuLeuGlnSerLeuLeuGlnMet 214
DB      126 GCCGTGATGGGGGTCAATCCGGGACCCGTGCTCTCCAGAGTCTGCTCAAGAAATG 185
QY      215 AsnThrLeuArgThrGlnGlyGlyHisLeuTrpAlaTrpAspTrpIleSerIleProLeu 234
DB      186 AACACCTTGCGTACTGATGATGAGACCTGTATGACCTTACGACCTGATGATGATGATGATG 245
QY      235 ValTrpThrGlnValValThrValAlaValIleSerPhePheLeuThrCysLeuValGly 254
DB      246 GTGTACACTCAGGTGTGATGACCGTGGCGGTATACAGCTTCTCTGCTGTGTGTGG 305
QY      255 ArgGlnPheLeuAsnProAlaLysAlaTrpProGlnHisGlnLeuAspLeuValValPro 274
DB      306 AGGCAATTTCTGAACCAACCAAGGCTTACCCGGGCAATGATGATGATGATGATGATGATGATG 365
QY      275 ValPheThrPheLeuGlnPhePhePheTrpValGlyTrpLeuValValAlaGlnLeu 294
DB      366 CTCTTCAGCTTCTGTGAGTTCTTCTTCTTATGCGCGGTGGCTGAGAGTGGCAGAGACGCTC 425
QY      295 IleAsnProPheGlnGlnAspAspAspPheGlnTrpAsnTrpIleValAspArgAsn 314
DB      426 ATCAACCCCTTTGGAGAGATGATGACACTTTGAGACCAACTGATGATGATGATGATGATGATG 485
QY      315 LeuGlnValSerLeuLeuAlaValAspGlnMetHisGlnAspLeuProArgMetGlnPro 334
DB      486 TTGCAGGTGTCTCTTCTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 545
QY      335 AspMetTrpTrpAsnLysProGlnProGlnProTrpTrpAlaAlaSerAlaGlnPhe 354
DB      546 GATATGTACTGGAATGATCCAGAACCAATCCCCCTTACACAGCTGCTTCTGCCAGTCT 605
QY      355 ArgArgAlaSerPheMetGlySerThrPheAsnIleSerLeuAsnLysGlnGlnMetGln 374
DB      606 CGCCGACCTCTCTTTTCTGCTCCACCTTCAATATCACTGCGGTAGAGAGACATGAG 665
QY      375 PheGlnProAsnGlnGlnAspGlnGlnAspAlaHisAlaGlyIleIleGlyArgPheLeu 394
DB      666 TTCCAGCGG-----GAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 719
QY      395 GlyLeuGlnSerHisAspHisHisProProArgAlaAsnSerArgThrLysLeuLeuTrp 414
DB      720 GGGCTGCAATCCACGACCAACAGCCCCCAGGCAAACTCAAGAGAACTCCTGTGTG 779
QY      415 ProLysArgGlnSerLeuLeuHisGlnGlyLeuProLysAsnHisLysAlaAlaLysGln 434
DB      780 CCCAAGAAAGAGAGCCATTTCCAGAGAGGCGCACCCCAAGAACTCAGGGGGGCCAGGCTG 839
QY      435 AsnValArgGlnGlnLysAspAsnLysAlaTrpLysLeuLysAlaValAspAlaPheLys 454
DB      840 GACTCTAGCAGACAGAGAGACAGAGAGGCTGAGG-----GAGGATGCTTCAAG 890
QY      455 SerGlyProLeuTrpGlnArgProGlyTrpTrpSerAlaProGlnThrProLeuSerPro 474
DB      891 TCCGCTGCTCTGTGTGGAGGCGGGGTTCACACAGCGCCCCACAGACACCTCGGCCAC 950
QY      475 ThrProMetPhePheProLeuGlnProSerAlaProSerLysLeuHisSerValThrGly 494
DB      951 ACCCTATGGTCTTCCACCCGAAAGATGAGCGCCCTTAGGTCTTGGCAGAGTCTGGAGC 1010

```

```

QY      495 IleAsp-----ThrLysAspLysSerLeuLysThrValSerSerGlyAlaLysLysSer 512
DB      1011 ATAGATGAGGCTGCCAAAGACAGAGGCTTACGCTTGCACCTCCAGACATCAAGAAAGAT 1070
QY      513 PheGlnLeuLeuSerGlnSerAspGlyAlaLeuMetGlnHisProGlnValSerGlnVal 532
DB      1071 TTCAGTGTCTCCAGAGAGGCGCGAGGCTCGGCGGAGCCCTCGCAAGAGGAGTCAAGT 1130
QY      533 ArgArgLysThrValGlnPheAsnLeuThrAspMetProGlnIleProGlnAsnHisLeu 552
DB      1131 AGAGAGAAAGAGCGTGGAGTTTAACCTGCGGATCTGTGAGAGGCCCCGAA--CATCTC 1187
QY      553 LysGlnProLeuGlnGlnSerProThrAsnIleHisThrThrLeuLysAspHisMetAsp 572
DB      1188 AAAAGAACCAATTGGAGCCAGCCAGTGGGATGACGATATCTCAAGACCAAGGAGAT 1247
QY      573 ProTrpTrpAlaLeuGlnAsnArgAspGlnAlaHisSer 585
DB      1248 CCTACTGGGCTTGGAAACAGGAGATGAAGCACTTCC 1286

```

Search completed: March 23, 2005, 15:22:03
 Job time : 7505.71 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2005, 22:00:09 ; Search time 10345.6 Seconds
(without alignments)
11376.592 Million cell updates/sec

Title: US-09-622-964a-4
Perfect score: 2429
Sequence: 1 CAGGAGTCCACCCAGCCTA.....AAAAAAAAAAAAAAAAAAAA 2429

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_to:*
11: gb_str:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2429	100.0	2429	6	BD136721
2	2418.4	99.6	2420	6	AF057170 Homo sapi
3	2013	82.9	2229	6	BD136720 Beest's ma
4	1995.4	82.1	2210	9	AF057169 Homo sapi
5	1953.2	80.4	2171	9	AF073501 Homo sapi
6	1942.8	80.0	2170	6	CQ729633 Sequence
7	1809.6	74.5	2187	9	AY357925 Macaca fa
8	1587	65.3	2441	9	BC015220 Homo sapi
9	1543.4	63.5	1758	9	AY515704 Homo sapi
10	1541.8	63.5	1758	6	AX745964 Sequence
11	1371	56.4	2435	6	BC041664 Homo sapi
12	1244.2	51.2	1263	6	BD132659 Secreted
13	712.8	29.3	2035	10	BC079048 Ratcun no
14	680.4	28.0	1904	10	AY450427 Mus muscu
15	658.2	27.1	1916	6	BD136744 Beest's ma
16	648.4	26.7	1289	4	AY064707 Sub scrof
17	642.2	26.4	16125	6	BD136719 Beest's ma
18	642.2	26.4	142092	9	AF139813 Homo sapi
19	642.2	26.4	163024	9	AP006260 Homo sapi

C	20	642.2	26.4	163915	2	AC087451	AC087451 Homo sapi
	21	642.2	26.4	166867	9	AP003733	AP003733 Homo sapi
	22	642.2	26.4	196080	0	AC004228	AC004228 Homo sapi
	23	640.6	26.4	706	9	HSVMD2P10	AF073499 Homo sapi
	24	640.6	26.4	112309	9	AC003025	AC003025 Human chr
	25	640.6	26.4	133683	2	AC084857	AC084857 Homo sapi
	26	599	24.7	726	9	AY357926S10	AY357935 Macaca fa
C	27	552.8	22.8	160169	2	AC051664	AC051664 Homo sapi
	28	431.4	17.8	2137	6	AX745966	AX745966 Sequence
	29	431	17.7	1530	6	CO731444	CO731444 Sequence
	30	431	17.7	1530	6	AX745993	AX745993 Sequence
	31	431	17.7	1530	9	AY515705	AY515705 Homo sapi
	32	431	17.7	1908	9	AF440756	AF440756 Homo sapi
	33	415	17.1	1954	10	AY450428	AY450428 Mus muscu
	34	415	17.1	1956	10	BC031186	BC031186 Mus muscu
	35	415	17.1	1957	10	BC019528	BC019528 Mus muscu
	36	415	17.1	1992	10	BC036157	BC036157 Mus muscu
	37	389.2	16.0	1861	9	AK000139	BC036163 Mus muscu
	38	369.4	15.2	1898	5	BC043854	AK000139 Homo sapi
	39	369.4	15.2	1921	5	AY273825	BC043854 Xenopus l
	40	365.2	15.0	1882	5	CR760914	AY273825 Xenopus l
	41	365.2	15.0	2021	5	BC061379	CR760914 Xenopus t
	42	358.4	14.8	492	9	HSVMD2P11	BC061379 Xenopus t
C	43	358.4	14.8	106648	2	AC004588	AF073500 Homo sapi
	44	358.2	14.7	2064	5	BC084429	AC004588 Homo sapi
	45	358.2	14.7	2064	5	BC084429	BC084429 Xenopus l

ALIGNMENTS

RESULT 1
LOCUS BD136721 2429 bp DNA linear PAT 18-SEP-2002
DEFINITION Beest's macular dystrophy gene.
BD136721
ACCESSION BD136721.1 GI:23231666
VERSION JP 2002504559-A/3.
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 (bases 1 to 2429)
Petrushkin,K., Caskey,T.C., Metzker,M. and Wadelius,C.
TITLE Beest's macular dystrophy gene
JOURNAL Patent: JP 2002504559-A 3 12-FEB-2002;
COMMENT MERCK & CO INC,CLAS WADLIUS
OS Homo sapiens (human)
PN JP 2002504559-A/3
PD 12-FEB-2002
PR 22-FEB-1999 JP 2000533447
PR 25-FEB-1998 US 60/075941,18-DEC-1998 US 60/112926 PI
WADLIUS
PC C07K16/18, C07K14/47, C12N5/10, C12N15/09, C12P19/34, C12Q1/68// PC
C12P21/08,
PC C12N5/00, C12N15/00
CC Beest's macular dystrophy gene
FH key location/Qualifiers
FT source 1..2429
FT location/Qualifiers
FT /organism="Homo sapiens (human)".

FEATURES

source

ORIGIN

Query Match 100.0%; Score 2429; DB 6; Length 2429;
Beest Local Similarity 100.0%; Pred. No. 0;
Matches 2429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 CAGGAGTCCACCCAGCCTTGTGTGATCATCGACCCACCTGGA 60

```
Db 1 CAGGAGTCCACAGCTAGTCCACAGCTTTCTGTGGATCATGGAACCCACCTGGAA 60
Qy 61 CCCCACTGACCCCAAGCCCACTGTGCAAGCCCACTGCTGACATGACCATCATCTTACA 120
Db 61 CCCCACTGACCCCAAGCCCACTGTGCAAGCCCACTGCTGACATGACCATCATCTTACA 120
Qy 121 CAAGCAAGTGGCTAATAGCCGCTTAGCTCTCTCCGCTGCTGCTGTGCTGCTGCTGCTG 180
Db 121 CAAGCAAGTGGCTAATAGCCGCTTAGCTCTCTCCGCTGCTGCTGTGCTGCTGCTGCTG 180
Qy 181 GCAGCATCTACAAAGCTCTATATGCGAGTCTTAACTCTTCTGCTCTGCTACTACATCA 240
Db 181 GCAGCATCTACAAAGCTCTATATGCGAGTCTTAACTCTTCTGCTCTGCTACTACATCA 240
Qy 241 TCCGCTTATTTATAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
Db 241 TCCGCTTATTTATAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
Qy 301 CTCTGATTGCGACAGCTACATCAGCTCATCCCACTTCTCTGCTGCTGCTGCTGCTG 360
Db 301 CTCTGATTGCGACAGCTACATCAGCTCATCCCACTTCTCTGCTGCTGCTGCTGCTG 360
Qy 361 TGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
Db 361 TGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
Qy 421 TCATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Db 421 TCATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Qy 481 GCAAGCTCATCCGCTACAGCTCAACCTGGCAACGCTCATCTCTGCTGCTGCTGCTG 540
Db 481 GCAAGCTCATCCGCTACAGCTCAACCTGGCAACGCTCATCTCTGCTGCTGCTGCTG 540
Qy 541 CAGTCTACAGAGGCTTCCCAAGCCCAAGCACTGCTGCTGCTGCTGCTGCTGCTGCTG 600
Db 541 CAGTCTACAGAGGCTTCCCAAGCCCAAGCACTGCTGCTGCTGCTGCTGCTGCTGCTG 600
Qy 601 CAGAACACAGAGCTTGGAGAACTGAGCTTACACACACATGTTCTGGGCTGCTGCTG 660
Db 601 CAGAACACAGAGCTTGGAGAACTGAGCTTACACACACATGTTCTGGGCTGCTGCTG 660
Qy 661 TGTGCTTGGCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Db 661 TGTGCTTGGCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Qy 721 TGTCTCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
Db 721 TGTCTCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
Qy 781 CCTACGACTGATTAATATCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
Db 781 CCTACGACTGATTAATATCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
Qy 841 GCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Db 841 GCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Qy 901 GCATGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Db 901 GCATGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Qy 961 GCTGCTGAGAGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
Db 961 GCTGCTGAGAGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
Qy 1021 ATGGCCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
Db 1021 ATGGCCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
Qy 1081 AAAGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
Db 1081 AAAGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
Qy 1141 GAGAAATGCTTGAACCCGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
Db 1141 GAGAAATGCTTGAACCCGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
Qy 1201 TGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1260
Db 1201 TGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1260
Qy 1261 CTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
Db 1261 CTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
Qy 1321 GAATAGCCGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380
Db 1321 GAATAGCCGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380
Qy 1381 CTTTATGAGCTTCACTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440
Db 1381 CTTTATGAGCTTCACTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440
Qy 1441 TCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
Db 1441 TCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
Qy 1501 CCAATGATACCAATCTTCCAGAGGCAAACTCAAGACCAAACTACTGCTGCTGCTGCTG 1560
Db 1501 CCAATGATACCAATCTTCCAGAGGCAAACTCAAGACCAAACTACTGCTGCTGCTGCTG 1560
Qy 1561 ATCCCTTCTCAGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620
Db 1561 ATCCCTTCTCAGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620
Qy 1621 CCAAGAGAGCAAAAGAGCTTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1680
Db 1621 CCAAGAGAGCAAAAGAGCTTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1680
Qy 1681 GTATCAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740
Db 1681 GTATCAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740
Qy 1741 CTTTCCCTTGAAGACATCAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800
Db 1741 CTTTCCCTTGAAGACATCAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800
Qy 1801 AAGCAAAAGCTTAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1860
Db 1801 AAGCAAAAGCTTAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1860
Qy 1861 GAGCGATGGGCTTGAATGAGAGCCCAAGAGCTATCTCAAGTGAAGAGAGAGAGAGAG 1920
Db 1861 GAGCGATGGGCTTGAATGAGAGCCCAAGAGCTATCTCAAGTGAAGAGAGAGAGAGAG 1920
Qy 1921 GTTTAACTGACGAGATATGAGAGATCCCGAAATCACTCAAAAGACCTTTGGAGACA 1980
Db 1921 GTTTAACTGACGAGATATGAGAGATCCCGAAATCACTCAAAAGACCTTTGGAGACA 1980
Qy 1981 ATCAGCAACCAACATACACTTCACTCAAAAGATCAATGATCTTATTTGGGCTTGA 2040
Db 1981 ATCAGCAACCAACATACACTTCACTCAAAAGATCAATGATCTTATTTGGGCTTGA 2040
Qy 2041 AAACAGGAGTAAACCACTTCTAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
Db 2041 AAACAGGAGTAAACCACTTCTAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
Qy 2101 CCTGACCTGTGTGTACACAGAGAGCACTGATCTGATCTGATCTGATCTGATCTGATCTG 2160
Db 2101 CCTGACCTGTGTGTACACAGAGAGCACTGATCTGATCTGATCTGATCTGATCTGATCTG 2160
Qy 2161 CTGAAGAACGCTGCTTCAACAGCTGATCAATATGCTTAAATGATTAATAAATCC 2220
Db 2161 CTGAAGAACGCTGCTTCAACAGCTGATCAATATGCTTAAATGATTAATAAATCC 2220
```

Qy 2221 AGACTACTTCAGCCTTTAATGCTTTTATTCATATAAACTGGAAGCTAGTGAACCA 2280
Db 2221 AGACTACTTCAGCCTTTAATGCTTTTATTCATATAAACTGGAAGCTAGTGAACCA 2280
Qy 2281 TTGGAACATTAACTCAGACTCTGAGATTCAAGCTGGGAACCTTACTTATCTGAAT 2340
Db 2281 TTGGAACATTAACTCAGACTCTGAGATTCAAGCTGGGAACCTTACTTATCTGAAT 2340
Qy 2341 CCAAGACAGCCAGCCTTATGATTAAGTCCCAACTTAAGATTAAATCAAAATCT 2400
Db 2341 CCAAGACAGCCAGCCTTATGATTAAGTCCCAACTTAAGATTAAATCAAAATCT 2400
Qy 2401 CGTTAAAAAATTT 2429
Db 2401 CGTTAAAAAATTT 2429

RESULT 2
AF057170 2420 bp mRNA linear PRI 17-OCT-1998
LOCUS Homo sapiens beetrophin (VMD2) mRNA, alternatively spliced product,
complete cds.
AF057170
AF057170.1 GI:3335160

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 2420)
Petrushkin, K., Koist, M. J., Bakall, B., Li, W., Xie, G., Marknell, T.,
Sengden, O., Forsman, K., Holmgren, G., Andreasson, S., Vujic, M.,
Berger, A. A., McCarly-Dugan, V., Figueroa, D., Austin, C. P.,
Metzger, M. L., Caskey, C. T. and Wadellius, C.
Identification of the gene responsible for Beet macular dystrophy
Nat. Genet. 19 (3), 241-247 (1998)
98324772
MEDLINE
9662395
2 (bases 1 to 2420)
Petrushkin, K.
Direct Submission
Submitted (03-APR-1998) Human Genetics, Merck Research
Laboratories, West Point, PA 19486, USA

FEATURES
source
1..2420
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q13"
1..2420
/gene="VMD2"
105..1412
/note="VMD2"
/note="Beet's macular dystrophy gene; BMD; alternatively
splice isoform; contains exon 7"
/product="beetrophin"
/protein_id="AAC64344.1"
/db_xref="GI:3335161"
/translation="MTITYSOVANBARLGSFRLLCMRGIVLYLGEPLIPILCY
IIRIVYALTEBOQLMFELTLCDSYIQLIPISFVLGRTVTVTRMMROYENLW
PDRMLVSGFVEBKDEQRLRLRLIRYALGNVLLRSVSTAVYRFSAOHLVQA
GFMTPAHKLKLEKSLPHNMFWVFWFANLSKRAWIGR.RDILLOSILNEKNTLR
TOGHLAYADWISLPLVYQVTVAVVSFLTCLVGRQFLIPAKAYGHELDVVPYF
TFLQPFYVGMILKVLGRALGMRHGGHGGHQLLETRMOCQERKVSREVSDAMWT
PVIPTAEAGSGLBGRRLIMWSSSTPLBERMMILIRPTGISTGICRCPCLMWR
CTRCGLMSKRTCTGISPSHPFTLLPSSVPEPLMAPPSISA"

ORIGIN
Query Match 99.6%; Score 2418.4; DB 9; Length 2420;
Beet local Similarity 100.0%; Pred. No. 0;

Matches 2419; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 CAGGAGGCCACAGGCTAGTGGCAAGACTTCTGTGGGATCATCGAACCCACTGGAA 60
Db 1 CAGGAGGCCACAGGCTAGTGGCAAGACTTCTGTGGGATCATCGAACCCACTGGAA 60
Qy 61 CCCACCTGACCCCAAGCCCACTGCTGAGCCCACTGCTGGCAGTACATCACTTACA 120
Db 61 CCCACCTGACCCCAAGCCCACTGCTGAGCCCACTGCTGGCAGTACATCACTTACA 120
Qy 121 CAAGCCAAATGCTAATGCCCCGCTTAAAGCTCTCTTCCCGCTGCTGTGTGGCGGG 180
Db 121 CAAGCCAAATGCTAATGCCCCGCTTAAAGCTCTCTTCCCGCTGCTGTGTGGCGGG 180
Qy 181 GCACATCTCAACATGCTATATGGCGAGTCTTAATCTTCTGCTGCTACTACATCA 240
Db 181 GCACATCTCAACATGCTATATGGCGAGTCTTAATCTTCTGCTGCTACTACATCA 240
Qy 241 TCCGCTTATTTATAGCTGGCCCTCACGGAAGAACACAGCTGATTTGAGAACTGA 300
Db 241 TCCGCTTATTTATAGCTGGCCCTCACGGAAGAACACAGCTGATTTGAGAACTGA 300
Qy 301 CTCTGATTTGAGCAGCTACATCCAGCTCATCCCATTTCTTGTGCTGGGCTTTACG 360
Db 301 CTCTGATTTGAGCAGCTACATCCAGCTCATCCCATTTCTTGTGCTGGGCTTTACG 360
Qy 361 TGACGCTGCTGCTGACCCCGCTGTGGAAACAGTACGGAACCTGCGGCCGACCGCC 420
Db 361 TGACGCTGCTGCTGACCCCGCTGTGGAAACAGTACGGAACCTGCGGCCGACCGCC 420
Qy 421 TCATGACCTGCTGCTGCGGCTCTGCTGAAAGGCAAGCAAGCCGCTGCTGCGGC 480
Db 421 TCATGACCTGCTGCTGCGGCTCTGCTGAAAGGCAAGCAAGCCGCTGCTGCGGC 480
Qy 481 GCACGCTCATCTGCTGCTGACCAACTGCGCAACGCTCATCTGCTGCGACGCTACACCG 540
Db 481 GCACGCTCATCTGCTGCTGACCAACTGCGCAACGCTCATCTGCTGCGACGCTACACCG 540
Qy 541 CAGCTTCAAGCCGCTTCCCAAGCCGCAAGCCGCTGCTGCAAGGCTTTATGACTCGCG 600
Db 541 CAGCTTCAAGCCGCTTCCCAAGCCGCAAGCCGCTGCTGCAAGGCTTTATGACTCGCG 600
Qy 601 CAGAACCAAGCAGTGTGGAAGAACTGAGCTACACACATGTTCTGGGTGCTGGG 660
Db 601 CAGAACCAAGCAGTGTGGAAGAACTGAGCTACACACATGTTCTGGGTGCTGGG 660
Qy 661 TGTGTTTGGCAACTGTCTCAATGAAGCGTGGCTTGGAGGTGCAATCCGGGACCTATCC 720
Db 661 TGTGTTTGGCAACTGTCTCAATGAAGCGTGGCTTGGAGGTGCAATCCGGGACCTATCC 720
Qy 721 TGTCTCAGAGCTGCTGTAAGAGATGAACACTTGGCTGCTCAATGTGAGACCTGTATG 780
Db 721 TGTCTCAGAGCTGCTGTAAGAGATGAACACTTGGCTGCTCAATGTGAGACCTGTATG 780
Qy 781 CTTACGACTGGAATTAATATCCCATCTGTTGATACACAGGTGAGTCTGTGGCGGTGACA 840
Db 781 CTTACGACTGGAATTAATATCCCATCTGTTGATACACAGGTGAGTCTGTGGCGGTGACA 840
Qy 841 GCTTCTTCTGACTTGTCTAATTTGGGCGGAGTTTGAACCCAGCCCAAGGCTTACCTG 900
Db 841 GCTTCTTCTGACTTGTCTAATTTGGGCGGAGTTTGAACCCAGCCCAAGGCTTACCTG 900
Qy 901 GCAATAGCTGAGCTCTGTTGTGCGGCTTTCAAGCTTCTGCAATTTCTTCTATATG 960
Db 901 GCAATAGCTGAGCTCTGTTGTGCGGCTTTCAAGCTTCTGCAATTTCTTCTATATG 960
Qy 961 GCTGCTGGAAGTGGGCTCTCCAGGCGCTGCTGAGGCTGAGGCAAGGAGGCTC 1020
Db 961 GCTGCTGGAAGTGGGCTCTCCAGGCGCTGCTGAGGCTGAGGCAAGGAGGCTC 1020
Qy 1021 ATGGCCAGCAGCTGCTTGAAGAGATGAGTGTACAGAAAGAAAGTCTTACGGGTAG 1080
Db 1021 ATGGCCAGCAGCTGCTTGAAGAGATGAGTGTACAGAAAGAAAGTCTTACGGGTAG 1080

QY 1081 AAAGCAGCCAGCGCTGTGTGCGCACACCTGTAATCCAGACTACGCGGAGGCTGAGGCGAG 1140
 DB 1081 AAAGCAGCGAGCGCTGTGTGCGCACACCTGTAATCCAGACTACGCGGAGGCTGAGGCGAG 1140
 QY 1141 GAGAAATGGCTTGAACCCGCGGAGCGGAGGTTGTGTGCGAGACAGCTCAATCAACCCCTT 1200
 DB 1141 GAGAAATGGCTTGAACCCGCGGAGCGGAGGTTGTGTGCGAGACAGCTCAATCAACCCCTT 1200
 QY 1201 TGGAGAGGATGATGATGATTTTGAACCAACTGGAATTGTGAGACAGAAATTTGAGAGTTC 1260
 DB 1201 TGGAGAGGATGATGATGATTTTGAACCAACTGGAATTGTGAGACAGAAATTTGAGAGTTC 1260
 QY 1261 CTTGTTGGCTGTGTGATGATGACACAGGACTGTGCTCGGATGAGACCGGACATGTAATG 1320
 DB 1261 CTTGTTGGCTGTGTGATGATGACACAGGACTGTGCTCGGATGAGACCGGACATGTAATG 1320
 QY 1321 GAATTAAGCCGAGCCGACAGCCGCTTACAGAGTGTCTTCCGCTGAGAGCTTC 1380
 DB 1321 GAATTAAGCCGAGCCGACAGCCGCTTACAGAGTGTCTTCCGCTGAGAGCTTC 1380
 QY 1381 CTTTATGAGGCTTCCAGCTTCAACATCAGCTGAAACAAAGAGATGAGTTTCAAGCCAA 1440
 DB 1381 CTTTATGAGGCTTCCAGCTTCAACATCAGCTGAAACAAAGAGATGAGTTTCAAGCCAA 1440
 QY 1441 TCAGAGGAGACGAGAGAGATGCTCAAGCTGAGCATGATGAGCTTCTTGAAGGCTGAGTC 1500
 DB 1441 TCAGAGGAGACGAGAGAGATGCTCAAGCTGAGCATGATGAGCTTCTTGAAGGCTGAGTC 1500
 QY 1501 CCATGATCAACATCTCTCCAGGAGCAAACTCAAGAGACCAACTGAGTGGCCCAAGAGGA 1560
 DB 1501 CCATGATCAACATCTCTCCAGGAGCAAACTCAAGAGACCAACTGAGTGGCCCAAGAGGA 1560
 QY 1561 ATCCCTTCTTCCAGAGGAGCTTCCCAAAAAACCAAGGAGGCAAGGCAAGGCTTGAAGG 1620
 DB 1561 ATCCCTTCTTCCAGAGGAGCTTCCCAAAAAACCAAGGAGGCAAGGCAAGGCTTGAAGG 1620
 QY 1621 CCAGAGAAACAAAGAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGG 1680
 DB 1621 CCAGAGAAACAAAGAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGG 1680
 QY 1681 GTATCAGAGGAGGCTTACTACAGTGGCCCAAGAGGCTTCAAGGCTTCAAGGCTTCAAG 1740
 DB 1681 GTATCAGAGGAGGCTTACTACAGTGGCCCAAGAGGCTTCAAGGCTTCAAGGCTTCAAG 1740
 QY 1741 GTATCAGAGGAGGCTTACTACAGTGGCCCAAGAGGCTTCAAGGCTTCAAGGCTTCAAG 1800
 DB 1741 GTATCAGAGGAGGCTTACTACAGTGGCCCAAGAGGCTTCAAGGCTTCAAGGCTTCAAG 1800
 QY 1801 AGACAAAGCTTAAAGCTGTGAGTTCGAGGAGGCAAGAAAGTTTGAATTTGCTTCAAG 1860
 DB 1801 AGACAAAGCTTAAAGCTGTGAGTTCGAGGAGGCAAGAAAGTTTGAATTTGCTTCAAG 1860
 QY 1861 GAGGATGGGAGCTTGTGAGAGCAACCAAGAGTATCTCAAGTGAAGGAGAAACTGTGGA 1920
 DB 1861 GAGGATGGGAGCTTGTGAGAGCAACCAAGAGTATCTCAAGTGAAGGAGAAACTGTGGA 1920
 QY 1921 GTTTAACTGAGGATATGAGAGATCCCGGAAATCACTCTCAAGAACTTTGGAACA 1980
 DB 1921 GTTTAACTGAGGATATGAGAGATCCCGGAAATCACTCTCAAGAACTTTGGAACA 1980
 QY 1981 ATCAACCAACCAACATACACTCACTCAAGATCATGAGATCTTAAATTTGGGCTTGA 2040
 DB 1981 ATCAACCAACCAACATACACTCACTCAAGATCATGAGATCTTAAATTTGGGCTTGA 2040
 QY 2041 AAAAGAGGATGAGACATTTCTTAAGCTGCTTCTTAATGAGGATGCTTGGACAGAGGT 2100
 DB 2041 AAAAGAGGATGAGACATTTCTTAAGCTGCTTCTTAATGAGGATGCTTGGACAGAGGT 2100
 QY 2101 CCTCACTGTGTGTACACAGAGAGACATGATCAAGCTCAAGGCAATACAGCTGTTCACA 2160
 DB 2101 CCTCACTGTGTGTACACAGAGAGACATGATCAAGCTCAAGGCAATACAGCTGTTCACA 2160

QY 2161 CTGAAGAACTGTGTCTTCAACAGAGCTTGAATCAAAATGTTAGCTTAAATGATTAATATCCC 2220
 DB 2161 CTGAAGAACTGTGTCTTCAACAGAGCTTGAATCAAAATGTTAGCTTAAATGATTAATATCCC 2220
 QY 2221 AGACTACTTGAAGCTTAAATGCTTTTATTCATATAAAACCTGAAACCTAGACTGAACCA 2280
 DB 2221 AGACTACTTGAAGCTTAAATGCTTTTATTCATATAAAACCTGAAACCTAGACTGAACCA 2280
 QY 2281 TTGGAACATTTAACTCAGACTGTGATTCAGAGTCCGGAACCTTATGTTCTATCTGAAT 2340
 DB 2281 TTGGAACATTTAACTCAGACTGTGATTCAGAGTCCGGAACCTTATGTTCTATCTGAAT 2340
 QY 2341 CCAAGACAGCCACCTTATGATTTCTGCGCAACTAATGATTTAATTAATCAAAATACT 2400
 DB 2341 CCAAGACAGCCACCTTATGATTTCTGCGCAACTAATGATTTAATTAATCAAAATACT 2400
 QY 2401 CGTTAAAAAATTTTTTTTTT 2420
 DB 2401 CGTTAAAAAATTTTTTTTTT 2420
 RESULT 3
 LOCUS BD136720 2229 bp DNA linear PAT 18-SEP-2002
 DEFINITION Best's macular dystrophy gene.
 ACCESSION BD136720
 VERSION BD136720.1 GI:23231665
 KEYWORDS JP 2002504559-A/2.
 SOURCE
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 AUTHORS Petrukhin, K., Caskey, T. C., Metzker, M. and Wadelius, C.
 TITLE Best's macular dystrophy gene
 JOURNAL Patent: JP 2002504559-A 2 12-FEB-2002;
 COMMENT
 OS Homo sapiens (human)
 PN JP 2002504559-A/2
 PD 12-FEB-2002
 PR 22-FEB-1999 JP 2000533447
 PR 25-FEB-1998 US 60/075941, 18-DEC-1998 US 60/112926 PI
 KONSTANTIN PETRUKHIN, THOMAS C CASKEY, MICHAEL METZKER, CLAES PI
 WADLIUS
 PC C07K16/18, C07K14/47, C12N5/10, C12N15/09, C12P19/34, C12Q1/68// PC
 C12P21/08,
 CC C12N5/00, C12N15/00
 CC Best's macular dystrophy gene
 FH Key Location/Qualifiers
 FT source 1..2229
 FT /organism='Homo sapiens (human)'.
 FEATURES
 source 1..2229
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Query Match 82.9%; Score 2013; DB 6; Length 2229;
 Best Local Similarity 91.6%; Pred. No. 0;
 Matches 226; Conservative 0; Mismatches 0; Indels 203; Gaps 1;
 QY 1 CAGGAGTCCAGCAGCTTGTGCGCAGACCTTGTGTGAGTCAATCGGACCCAGCTTGA 60
 DB 1 CAGGAGTCCAGCAGCTTGTGCGCAGACCTTGTGTGAGTCAATCGGACCCAGCTTGA 60
 QY 61 CCCACCTGAGCCCAAGCCCACTGCTGAGGCCCACTGCTGGCCATGACATCACTTACA 120
 DB 61 CCCACCTGAGCCCAAGCCCACTGCTGAGGCCCACTGCTGGCCATGACATCACTTACA 120
 QY 121 CAAACCAAGTGTGATATGAGCTTGTGAGCTTGTGAGCTTGTGAGCTTGTGAGCTTGTGAG 180
 DB 121 CAAACCAAGTGTGATATGAGCTTGTGAGCTTGTGAGCTTGTGAGCTTGTGAGCTTGTGAG 180

QY 181 GCAGACTCTACAAAGCTGTATATGCGAGATTCTTAATCTTCCTGCTGTACTACTACATCA 240
DB 181 GCAGACTCTACAAAGCTGTATATGCGAGATTCTTAATCTTCCTGCTGTACTACTACATCA 240
QY 241 TCCGCTATTATTTATTAAGCTGTGCGCTTCCACAGAAAGAAACAAGCTGATGTTTGAAAACTGA 300
DB 241 TCCGCTATTATTTATTAAGCTGTGCGCTTCCACAGAAAGAAACAAGCTGATGTTTGAAAACTGA 300
QY 301 CTCTGTATTGGAACAGCTACATCCAGCTCATTCGCCATTTCCTGCTGTGGCTTCTACG 360
DB 301 CTCTGTATTGGAACAGCTACATCCAGCTCATTCGCCATTTCCTGCTGTGGCTTCTACG 360
QY 361 TGACGCTGTGCTGACCCGCTGTGTGAAACAGTACAGAAACCTTGCCGTGTGCCACGCC 420
DB 361 TGACGCTGTGCTGACCCGCTGTGTGAAACAGTACAGAAACCTTGCCGTGTGCCACGCC 420
QY 421 TCATGAGCTGTGTGTGCGGCTTTCGTGAAGGCAAGCAAGCAAGGCCCTGTGCTGCGGC 480
DB 421 TCATGAGCTGTGTGTGCGGCTTTCGTGAAGGCAAGCAAGCAAGGCCCTGTGCTGCGGC 480
QY 481 GCAGGCTATCCGCTACAGCCAACTGTGGCAAGTGTCTATCTGTGGCAACGTCAAGACCG 540
DB 481 GCAGGCTATCCGCTACAGCCAACTGTGGCAAGTGTCTATCTGTGGCAACGTCAAGACCG 540
QY 541 CAGTCTACAAAGGCTTCCCAAGCGCCCAAGCACTGTGTCAAGCAGGCTTTATGACTCGG 600
DB 541 CAGTCTACAAAGGCTTCCCAAGCGCCCAAGCACTGTGTGTCAAGCAGGCTTTATGACTCGG 600
QY 601 CAGAACACAAAGCTGTGAGAACTGAGCTTACCAACAACATGTTCTTGTGCTGCTGGG 660
DB 601 CAGAACACAAAGCTGTGAGAACTGAGCTTACCAACAACATGTTCTTGTGCTGCTGGG 660
QY 661 TGTGTTTGGCCAACTGTCAATGAAGGCTGTGGAGTGTGAATCCGAGACCTTATCC 720
DB 661 TGTGTTTGGCCAACTGTCAATGAAGGCTGTGGAGTGTGAATCCGAGACCTTATCC 720
QY 721 TGTCTCAGAGCTGTGAAACGAGATGAACACCTTGCGTACTCAGTGTGAGCAACCTGTATG 780
DB 721 TGTCTCAGAGCTGTGAAACGAGATGAACACCTTGCGTACTCAGTGTGAGCAACCTGTATG 780
QY 781 CCTACGACTGTGATTAATCCACTGTGTATATACACAGGTGTGATCTGTGTGCGGTGTACA 840
DB 781 CCTACGACTGTGATTAATCCACTGTGTATATACACAGGTGTGATCTGTGTGCGGTGTACA 840
QY 841 GCTTCTTCTGACTGTGTCTAAGTGTGGCGGCAAGTTCTGAACCCAGCCCAAGCCTTACCTG 900
DB 841 GCTTCTTCTGACTGTGTCTAAGTGTGGCGGCAAGTTCTGAACCCAGCCCAAGCCTTACCTG 900
QY 901 GCATGAGCTGGAACCTGTGTGCGGCTTCAAGGCTTCAAGTGTCTTCTTCTATGTTG 960
DB 901 GCATGAGCTGGAACCTGTGTGCGGCTTCAAGGCTTCAAGTGTCTTCTTCTATGTTG 960
QY 961 GCTGACTGAAAGGTGGGCTCTTCCAGGGGCCCTGTGGGCTGTGAGGATGCCAGAGGGTTC 1020
DB 961 GCTGACTGAA----- 970
QY 1021 ATGGCCAGACAGCTGTTGAGACAGAGATGCAAGTGTCAAGAAAAGAAAGTCTTCAAGGGTAG 1080
DB 971 ----- 970
QY 1081 AAAGAGCCAGAGCGTGTGTGGGCAACCTGTAAATCCAGCTAATCCGAGGCTGAGGAG 1140
DB 971 ----- 970
QY 1141 GAGAAATCGCTTGAAACCCGAGAGCGAGAGTGTGTGTGAGAGCAGCTCATCAACCCCTT 1200
DB 971 ----- 970
QY 1201 TGAAGAGATGATGATGATTTTGAAGACCAACTGATGTGTGACAGAAATTTGCAAGGTTC 1260
DB 998 TGAAGAGATGATGATGATTTTGAAGACCAACTGATGTGTGACAGAAATTTGCAAGGTTC 1057

QY 1261 CCTGTGTGCTGTGATGATGATGACACAGAACCTGTCTGTGATGAGGCCGACATGTACTG 1320
DB 1058 CCTGTGTGCTGTGATGATGATGACACAGAACCTGTCTGTGATGAGGCCGACATGTACTG 1117
QY 1321 GAATTAAGCCGAGCCACAGCCCTTACACAGCTGTCTTCCGCCAGTTCCGTGAGCTTC 1380
DB 1118 GAATTAAGCCGAGCCACAGCCCTTACACAGCTGTCTTCCGCCAGTTCCGTGAGCTTC 1177
QY 1381 CTTTATGGGCTCCACTTCAACATTCAGCTGAACAAAGAGAGATGAGATTCACGCCAA 1440
DB 1178 CTTTATGGGCTCCACTTCAACATTCAGCTGAACAAAGAGAGATGAGATTCACGCCAA 1237
QY 1441 TCAGAGAGACAGAGAGATGCTACAGCTGTGATCATTTGCGCTTCTTAAGCCTGACAGTC 1500
DB 1238 TCAGAGAGACAGAGAGATGCTACAGCTGTGATCATTTGCGCTTCTTAAGCCTGACAGTC 1297
QY 1501 CCATGATCACCATCTCTCCAGAGGCAAACTCAAGAACCAACTACTGTGGCCCAAGAGGA 1560
DB 1298 CCATGATCACCATCTCTCCAGAGGCAAACTCAAGAACCAACTACTGTGGCCCAAGAGGA 1357
QY 1561 ATCCCTTCTCGACGAGGCGCTGCCCAAAAACACAGGCAAGCCAAACGTTAAGGGG 1620
DB 1358 ATCCCTTCTCGACGAGGCGCTGCCCAAAAACACAGGCAAGCCAAACGTTAAGGGG 1417
QY 1621 CCAGGAAGACAAAGGCTGTGAAGCTTAAAGCTGTGAGCCCTTCAAGTCTGCGCCACT 1680
DB 1418 CCAGGAAGACAAAGGCTGTGAAGCTTAAAGCTGTGAGCCCTTCAAGTCTGCGCCACT 1477
QY 1681 GTATCAGAGGCAAGGCTACTACAGTGTGCCCAAGAGCGCCCTCAGGCCACTGCCATGTT 1740
DB 1478 GTATCAGAGGCAAGGCTACTACAGTGTGCCCAAGAGCGCCCTCAGGCCACTGCCATGTT 1537
QY 1741 CTTCCCTTGAACATCAAGGCGCTGTCAAACTTCAAGTGTTCACAGGCAATGACCAACAA 1800
DB 1538 CTTCCCTTGAACATCAAGGCGCTGTCAAACTTCAAGTGTTCACAGGCAATGACCAACAA 1597
QY 1801 AGACAAAAGCTTAAAGACTGTGAGTCTGTGGGCCCAAGAAAGTTTGAATTGCTCTAGA 1860
DB 1598 AGACAAAAGCTTAAAGACTGTGAGTCTGTGGGCCCAAGAAAGTTTGAATTGCTCTAGA 1657
QY 1861 GAGCGATGGGGCTTGTATGAGAACCCAGAAAGTATCTCAAGTGAAGAGAGAAACTGTGGA 1920
DB 1658 GAGCGATGGGGCTTGTATGAGAACCCAGAAAGTATCTCAAGTGAAGAGAGAAACTGTGGA 1717
QY 1921 GTTTAACTGACGATATGACAGAGATCCCGGAAATACCTCAAGAAACCTTTGGAACA 1980
DB 1718 GTTTAACTGACGATATGACAGAGATCCCGGAAATACCTCAAGAAACCTTTGGAACA 1777
QY 1981 ATCAACAAACCAATPACACTACACTCAAGATCAATGATGATCTTATTTGGGCTTGGGA 2040
DB 1778 ATCAACAAACCAATPACACTACACTCAAGATCAATGATGATCTTATTTGGGCTTGGGA 1837
QY 2041 AAACAGGATGAAAGCAATTCCTAACCTGTCTCTAATGGGAGTGTTCGCAAGCAGGT 2100
DB 1838 AAACAGGATGAAAGCAATTCCTAACCTGTCTCTAATGGGAGTGTTCGCAAGCAGGT 1897
QY 2101 CCTCACTGTGTGTACACACAGCAGACACTGATTCAGTGTCAAGGCAATACACTGTGCCA 2160
DB 1898 CCTCACTGTGTGTACACACAGCAGACACTGATTCAGTGTCAAGGCAATACACTGTGCCA 1957
QY 2161 CTGAAGAAAGTGTCTCAACAGCCTGAAATCAATGTTAGCTTAAATGATPAAAAATCC 2220
DB 1958 CTGAAGAAAGTGTCTCAACAGCCTGAAATCAATGTTAGCTTAAATGATPAAAAATCC 2017
QY 2221 AGACTACTTCAGCCTTTAATGCTTTTAAATGATPAAAACTGTGAAGAGCTAGAACCA 2280
DB 2018 AGACTACTTCAGCCTTTAATGCTTTTAAATGATPAAAACTGTGAAGAGCTAGAACCA 2077
QY 2281 TTGAAACATTTAACTGACACTGTGATTCAGAGTCCGGAACCTTGTAGTTCTATCTGAAT 2340
DB 2078 TTGAAACATTTAACTGACACTGTGATTCAGAGTCCGGAACCTTGTAGTTCTATCTGAAT 2137
QY 2341 CCAAGACAGCCACCTTATGATGATGCTGCCAAACTAATGATTTAATPAAAAATPACT 2400

Db	2138	CCAAAGACAGCCACCTTACTGATATACCTGCCAAACTATAGATTATATATACAAATCT	2197
Qy	2401	CGTTAAAAA	2429
Db	2198	CGTTAAAAA	2226
RESULT 4	AF057169		
LOCUS	AF057169		
DEFINITION	Homo sapiens bestrophin (VMD2) mRNA, alternatively spliced product,		
ACCESSION	AF057169		
VERSION	AF057169.1		
KEYWORDS	GI:3335158		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
AUTHORS	1 (bases 1 to 2210) Petrushkin, K., Koishi, M.J., Bakall, B., Li, W., Xie, G., Maknelli, T., Sandgren, O., Foreman, K., Holmgren, G., Andreasson, S., Vujic, M., Bergen, A.A., McGarty, Dugan, V., Figueroa, D., Austin, C.P., Metzger, M.T., Caskey, C.T. and Wadelius, C. Identification of the gene responsible for Best macular dystrophy Nat. Genet. 19 (3), 241-247 (1998)		
TITLE	2 (bases 1 to 2210) Petrushkin, K.		
JOURNAL	Direct Submision		
REFERENCE	Submitted (03-APR-1998) Human Genetics, Merck Research Laboratories, West Point, PA 19486, USA		
AUTHORS	Location/Qualifiers		
TITLE	1.. 2210		
JOURNAL	/organism="Homo sapiens"		
REFERENCE	/mol_type="mRNA"		
AUTHORS	/db_xref="taxon:9606"		
TITLE	/chromosome="11"		
JOURNAL	/map="11q13"		
REFERENCE	1.. 2210		
AUTHORS	/gene="VMD2"		
TITLE	105.. 1862		
JOURNAL	/gene="VMD2"		
REFERENCE	/note="Best's macular dystrophy gene; BMD; alternatively spliced isoform"		
AUTHORS	/codon_start=1		
TITLE	/product="bestrophin"		
JOURNAL	/protein_id="AAC64343.1"		
REFERENCE	/db_xref="GI:3335158"		
AUTHORS	/translation="MTTITTSQVANNRLGSFSLILCMRGSIYKLLYGEFFILLCY IIRITRYALTEBQQLMEFKLTLYCDSYQLPISEVLGFTYTLVTRWNGYENLP PDRIMSLVSGVEKQDEQGLRLIRLIRYANLGNVLILSVSTAYAKPPSKOHLVQA GFMTPAHKQLEKLSLPHNMFWPWPWAFNLISKAWLGRIRDPILLOSILNEMWTL TOCGHIALDWMSIPLVYTOVTVAAVYSEFLTCITVGRPLNPAKYVPGHELDIVYV TDFWKKREPPQPYTAAQAOPRASAOPRASPMSGSTPNISLKNEMEFPQNEDEDAAGITIG PLIGQSIDHPRPRANSRTKLMPKRSLSLHEGLPNHKAQNVGQEDNKAWKIKAV DARSAFLYORPGYISAPOTPLSPTEWFPPLPSAPSKLSVGTGIDTQKSLKTVSSG AKSFELLSDDADLMEHPVSVQRKRTVEFNLTDPETLPEHNHLEKPEQSPTNHTT LKHMDPYMALENRDEAHS"		
ORIGIN			
Query Match	82.1%;	Score 1995.4;	DB 9;
Best Local Similarity	91.5%;	Prod. No. 0;	Length 2210;
Matches 2209;	Conservative 0;	Mismatches 1;	Indels 203;
			Gaps 1;
Qy	1	CAGGAGTCCACACAGCCTAGTGCAGACACTTCTGTGGATCATCGACCACTGGAA	60
Db	1	CAGGAGTCCACACAGCCTAGTGCAGACACTTCTGTGGATCATCGACCACTGGAA	60
	61	CCCCACTGACCAAGCCCACTGCTGCAGACCCCACTGCTGGCCATGACATCACTTACA	120

Db	61	CCCACTGACCCAGGCCCACTGCTGAGCCACTGCGCATGACCATCATTTACA	12
Qy	121	CAAGCCAAATGCTTAAATGCGGCTTAGGCTCTTCTCCGCGCTGCTGTGCTGCGG	18
Db	121	CAGGCCAAGTGTCTAATATCCGCTTAGCGCTCTTCTCCGCGCTGCTGTGCTGCGG	18
Qy	181	GGAGATCTCAAGAGCTGTATATGGACAGTTCTTAATCTTCCGCTGCTGCTAATCA	24
Db	181	GGAGATCTCAAGAGCTGTATATGGACAGTTCTTAATCTTCCGCTGCTGCTAATCA	24
Qy	241	TCCGCTTAATTAATAGCTGCGCTCTCACGGAAACAACAGCTGATGTTTGAACCTGA	30
Db	241	TCCGCTTAATTAATAGCTGCGCTCTCACGGAAACAACAGCTGATGTTTGAACCTGA	30
Qy	301	CTCTATTTTGGAGAGCTACATCCAGTCAATCCCCATTTCTTCCGCTGCGCTTACG	36
Db	301	CTCTATTTTGGAGAGCTACATCCAGTCAATCCCCATTTCTTCCGCTGCGCTTACG	36
Qy	361	TGACGCTGTGTCGACCCGCTGGTGGAAACAAGTACAGAACTGCGCGTGGCCGACG	42
Db	361	TGACGCTGTGTCGACCCGCTGGTGGAAACAAGTACAGAACTGCGCGTGGCCGACG	42
Qy	421	TGATGAGCTGTGTGCGGCTTGTCTGAAGCAAGACGAGCAGCGCTGTGCGGC	48
Db	421	TGATGAGCTGTGTGCGGCTTGTCTGAAGCAAGACGAGCAGCGCTGTGCGGC	48
Qy	481	GCAGGCTCATCCGCTACGCAACTGGGGCAAGTGTCTATCTTGGCGAGGCTGACACG	54
Db	481	GCAGGCTCATCCGCTACGCAACTGGGGCAAGTGTCTATCTTGGCGAGGCTGACACG	54
Qy	541	CAGTCTAACAAGCGCTTCCCGAGCGCCAGCACTGTGTCAAGCAGGCTTTATGACTCGG	60
Db	541	CAGTCTAACAAGCGCTTCCCGAGCGCCAGCACTGTGTCAAGCAGGCTTTATGACTCGG	60
Qy	601	CAGAAACAAGCAGTGGAGAACTGAGCTTACACACACAATGTTCTGGGTGCTCGG	66
Db	601	CAGAAACAAGCAGTGGAGAACTGAGCTTACACACACAATGTTCTGGGTGCTCGG	66
Qy	661	TGTGTTTGGCAACTGTCAATGAAAGCGGTGGCTTGGAGGTGCAATCCGGACCTTATCC	72
Db	661	TGTGTTTGGCAACTGTCAATGAAAGCGGTGGCTTGGAGGTGCAATCCGGACCTTATCC	72
Qy	721	TGCTCCAGAGCTGTGGAACGAGATGAACAACCTTGCCTCACTGATGGAACCTGTATG	78
Db	721	TGCTCCAGAGCTGTGGAACGAGATGAACAACCTTGCCTCACTGATGGAACCTGTATG	78
Qy	781	CTTACGACCTGAGTATATCCACTGTGTGTATACACAGGTGTGACTGTGGCGGTATAC	84
Db	781	CTTACGACCTGAGTATATCCACTGTGTGTATACACAGGTGTGACTGTGGCGGTATAC	84
Qy	841	GCTTCTTCCGCACTTGTCTAGTTGGGCGGAGTTTCTGAACCCAGCAAGGCTTACCGTG	90
Db	841	GCTTCTTCCGCACTTGTCTAGTTGGGCGGAGTTTCTGAACCCAGCAAGGCTTACCGTG	90
Qy	901	GCCATAGCTGGAACCTCGTGTGCGCGCTTCAAGTTCCTGAGTTCCTTCTATGTTG	96
Db	901	GCCATAGCTGGAACCTCGTGTGCGCGCTTCAAGTTCCTGAGTTCCTTCTATGTTG	96
Qy	961	GCTGCTGGAAGTGGGCTCTCCAGGCGCTGCTGGGCTGAGGCAATGGCCAGAGGCGTC	102
Db	961	GCTGCTGGAAGTGGGCTCTCCAGGCGCTGCTGGGCTGAGGCAATGGCCAGAGGCGTC	102
Qy	1021	ATGGCCAGAGCTGCTTGAGACGAGATGCAAGTGTGAGAAAGAAAGTCTCACGCGTAG	108
Db	1021	ATGGCCAGAGCTGCTTGAGACGAGATGCAAGTGTGAGAAAGAAAGTCTCACGCGTAG	108
Qy	1081	AAAGCAGCGAGCGTGTGCGCACACCTGTAAATCCCACTACTCGGAGGCTGAGCAG	114
Db	1081	AAAGCAGCGAGCGTGTGCGCACACCTGTAAATCCCACTACTCGGAGGCTGAGCAG	114
Qy	1141	GAGAAATGCTTGAACCGGGAGCGGAGGTTGTGTGGCAGAGCAGCTTATCAACCTT	120
Db	1141	GAGAAATGCTTGAACCGGGAGCGGAGGTTGTGTGGCAGAGCAGCTTATCAACCTT	120


```

Db      971 -----GATGGCAGAGCAGCTCATCAACCCCTT 997
Qy      1201 TTGAGAGGATGATGATGATTTTGGAGCAACCTGGATTGTGCAGAGGATTTGGAGGTGC 1260
Db      998 TGGAGAGGATGATGATGATTTTGGAGCAACCTGGATTGTGCAGAGGATTTGGAGGTGC 1057
Qy      1261 CTTGTTGGCTGTGTGATGATGATGACACAGAGCCTTCCTGATGATGAGCCGAGCATGTACTG 1320
Db      1058 CTTGTTGGCTGTGTGATGATGATGACACAGAGCCTTCCTGATGATGAGCCGAGCATGTACTG 1117
Qy      1321 GAATTAAGCCGAGGACACAGCCCTTACACAGCTGCTTCGCGCAAGTTCCGTGAGCCTC 1380
Db      1118 GAATTAAGCCGAGGACACAGCCCTTACACAGCTGCTTCGCGCAAGTTCCGTGAGCCTC 1177
Qy      1381 CTTTATGGGCTCCACTTCAACATCAGCTGTAACAAAGAGAGATGAGATTCGAGCCCA 1440
Db      1178 CTTTATGGGCTCCACTTCAACATCAGCTGTAACAAAGAGAGATGAGATTCGAGCCCA 1237
Qy      1441 TCAAGAGAGAGAGAGAGATGCTCAGCGTGGCATTTGGCCGCTTCTAGGCTGCAATC 1500
Db      1238 TCAAGAGAGAGAGAGAGATGCTCAGCGTGGCATTTGGCCGCTTCTAGGCTGCAATC 1297
Qy      1501 CCATGATCACCATTCTCCACAGGSCAAACTCAAGACCAAACTACTGTGKCCCAAGAGGGA 1560
Db      1298 CCATGATCACCATTCTCCACAGGSCAAACTCAAGACCAAACTACTGTGKCCCAAGAGGGA 1357
Qy      1561 ATCCCTTCTCAGAGGCGCTGCCCAAAAACAAGGCGAGCCAAACAGAACTTTAGGGG 1620
Db      1358 ATCCCTTCTCAGAGGCGCTGCCCAAAAACAAGGCGAGCCAAACAGAACTTTAGGGG 1417
Qy      1621 CCAAGAGAGACAAACAAGGCTGGAAGCTGAGGACGCTTCAAGTCTGAGCCCACT 1680
Db      1418 CCAAGAGAGACAAACAAGGCTGGAAGCTGAGGACGCTTCAAGTCTGAGCCCACT 1477
Qy      1681 GATCAGAGGCGAGCTACTACAGTGTCCCAAGACGCGCTTCAAGCCCACTCCCATGTT 1740
Db      1478 GATCAGAGGCGAGCTACTACAGTGTCCCAAGACGCGCTTCAAGCCCACTCCCATGTT 1537
Qy      1741 CTTCCCTTGAACCATCAGGCGCTGTAAGCTTCAAGTGTCAAGGCAATGAGACCA 1800
Db      1538 CTTCCCTTGAACCATCAGGCGCTGTAAGCTTCAAGTGTCAAGGCAATGAGACCA 1597
Qy      1801 AGACAAAGCTTAAAGACTGATGATGAGGCGCAAGAAAGTTTGAATTCCTCTGAGA 1860
Db      1598 AGACAAAGCTTAAAGACTGATGATGAGGCGCAAGAAAGTTTGAATTCCTCTGAGA 1657
Qy      1861 GAGCGATGGGCGCTTGAATGAGACCCAGAAAGTATCTCAAGTGAAGAGAAACTGTGGA 1920
Db      1658 GAGCGATGGGCGCTTGAATGAGACCCAGAAAGTATCTCAAGTGAAGAGAAACTGTGGA 1717
Qy      1921 GTTTAACTGACGGATATGCGAGAGATCCCGGAAATCACTTCAAGAGACCTTTGGAACA 1980
Db      1718 GTTTAACTGACGGATATGCGAGAGATCCCGGAAATCACTTCAAGAGACCTTTGGAACA 1777
Qy      1981 ATCAGCAACCAACTACACTCACTCAAAAGATCAGATGATGATGAGGCGCTTGA 2040
Db      1778 ATCAGCAACCAACTACACTCACTCAAAAGATCAGATGATGATGAGGCGCTTGA 1837
Qy      2041 AAACAGGAGTGAAGACATTCCTTAACCTGCTTCTTAATGGGGATGCTTCGCGACAGAGT 2100
Db      1838 AAACAGGAGTGAAGACATTCCTTAACCTGCTTCTTAATGGGGATGCTTCGCGACAGAGT 1897
Qy      2101 CCTTCACTGTGTGTATGACACAGAGAGACATGATCAGTCAACAGCATTCAGCTTCCACA 2160
Db      1898 CCTTCACTGTGTGTATGACACAGAGAGACATGATCAGTCAACAGCATTCAGCTTCCACA 1957
Qy      2161 CTGAAGAGAGCTGTCTCAACAGAGCTGAATCAAAATGGTTAGCTTAATAGATTAATCCC 2220
Db      1958 CTGAAGAGAGCTGTCTCAACAGAGCTGAATCAAAATGGTTAGCTTAATAGATTAATCCC 2017
Qy      2221 AGACTACTTACGCTTTAATGCTTTTATGATTAATAAACTGTGAAGCTAGACTGAACCA 2280
Db      2018 AGACTACTTACGCTTTAATGCTTTTATGATTAATAAACTGTGAAGCTAGACTGAACCA 2077

```

```

Qy      2281 TTGAAACATTTAACTCAGACTGTGATTCAGAGTCCGGAACCTTAGTTCATCTGAT 2340
Db      2078 TTGAAACATTTAACTCAGACTGTGATTCAGAGTCCGGAACCTTAGTTCATCTGAT 2137
Qy      2341 CCAAGACAGCCACCTTAGTATCTGCCCAAACTAATGATTTAATTAATCAAAATCT 2400
Db      2138 CCAAGACAGCCACCTTAGTATCTGCCCAAACTAATGATTTAATTAATCAAAATCT 2197
Qy      2401 CGTTAAAAAAA 2413
Db      2198 CGTTAAAAAAA 2210

```

RESULT 5

AF073501

LOCUS

Homo sapiens vitelliform macular dystrophy protein (VMD2) mRNA, PRI 02-SEP-1998

DEFINITION

ACCESSION

AF073501

VERSION

AF073501.1

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1..2171

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/chromosome="11"

/map="11q12-q13.1"

/issue_type="retinal pigment epithelium"

1..2171

/gene="VMD2"

89..1846

/gene="VMD2"

/note="Best disease"

/codon_start=1

/product="vitelliform macular dystrophy protein"

/protein_id="AAC33766.1"

/db_xref="GI:3511242"

/translation="MTITVTSQVAAARLGSFSSRLLCWRGSIYKLYGEPLIFLCY

IIRFIRLALTEBQDLMEFLTYCDYIOLIPFLGFTVTLVVRMWNQENLPM

PRLMSLVSGFVGEKDEGRLLRRTILRYANLGNVLIRSVSTAVYRPSAOHLVA

GPMTPRBNKOLEKLSLPHNMFVPMVFAFLSMKAWLGRIRIDILQSLINENYLR

TQCGHLYAVDWISILPLVYQVTVAVVSFELCLVGRQPLPAKAYGHELDLVVPV

TPIQAPFYVWMLYARGLINPRGDDDPFTNMTVPSNLQVSLIADNEMQDILPMPR

DMYMKRPEQPTTASAPRRASFKGSTNISLNSKEMEFQPNQDEBDNACTGR

FLIQSHDHNHPRPANSRTKLWPKRESLLHEGLPKHKAKAKONVGGEDKAWLKLAKV

DAFKSAPLYORPGVYSAPOPTLSPTPWFPLPSASPKLHSVTGIDTKDGLKTVSG

AKSFELISESDLMENREVSQVRKTVFNLTDMPBEIPENHLKEPLQESPTNIHTT

LKHMPDYWMLERDEAHS"

ORIGIN

Query Match 80.4%; Score 1953.2; DB 9; Length 2171;
 Best Local Similarity 91.3%; Pred. No. 0;
 Matches 2168; Conservative 0; Mismatches 3; Indels 203; Gaps 1;


```
OY 17 CCTAGTCCGCAAGACCTTCTGTGGGATATGGAACCCACCTGGAAACCCACCTGACCCAG 76
Db 1 CCTAGTCCGCAAGACCTTCTGTGGGATATGGAACCCACCTGGAAACCCACCTGACCCAG 60
OY 77 CCCACCTGTCTGACAGCCCACTGCTGTGGCCATGACCATTAACAAGCCAAAGTGGCTAA 136
Db 61 CCCACCTGTCTGACAGCCCACTGCTGTGGCCATGACCATTAACAAGCCAAAGTGGCTAA 120
OY 137 TGCCCGCTTAAAGCTCTCTCCGCGCTGTGTGTGTGGCGGGGACATTTACAAAGCT 196
Db 121 TGCCCGCTTAAAGCTCTCTCCGCGCTGTGTGTGTGGCGGGGACATTTACAAAGCT 180
OY 197 GCTATATGGGAGTTCTTAAATCTTCTGTCTGTCTACTACATCATCGCTTATTTATAG 256
Db 181 GCTATATGGGAGTTCTTAAATCTTCTGTCTGTCTACTACATCATCGCTTATTTATAG 240
OY 257 GCTGGCCCTCAACGGGAAGAAACAACAGCTGATGTTTGAAGAACTGACTCTGTATTTGCGACAG 316
Db 241 GCTGGCCCTCAACGGGAAGAAACAACAGCTGATGTTTGAAGAACTGACTCTGTATTTGCGACAG 300
OY 317 CTATATCAGACTATCCCATTTCTTCTGTGTGGGCTTCTAGTGAAGCTGTGTGTGAC 376
Db 301 CTATATCAGACTATCCCATTTCTTCTGTGTGGGCTTCTAGTGAAGCTGTGTGTGAC 360
OY 377 CCGCTGTGGAACCAAGTACGAGAACTGTGCGTGGCCGACCGCTCATGAGCCTGTGTGAC 436
Db 361 CCGCTGTGGAACCAAGTACGAGAACTGTGCGTGGCCGACCGCTCATGAGCCTGTGTGAC 420
OY 437 GGGCTTGTGGAAGGAGGAGGAGGAGGCGGCGTGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 496
Db 421 GGGCTTGTGGAAGGAGGAGGAGGCGGCGTGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480
OY 497 CGGCAACTGTGGGCAAGTGTCTATCTGTGGACCGTCAAGCAGCAGCAGTCTAACAAGCGCTT 556
Db 481 CGGCAACTGTGGGCAAGTGTCTATCTGTGGACCGTCAAGCAGCAGCAGTCTAACAAGCGCTT 540
OY 557 CCCCAAGCGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 616
Db 541 CCCCAAGCGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 600
OY 617 GGAAGAACTGAGCCTTACCAACAACATGTTCTGTGCGCTGTGGGCTGTGGGCTGTGGCAACT 676
Db 601 GGAAGAACTGAGCCTTACCAACAACATGTTCTGTGCGCTGTGGGCTGTGGGCTGTGGCAACT 660
OY 677 GTCAATGAAGCGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 736
Db 661 GTCAATGAAGCGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 720
OY 737 GAACGAGATGAACACTTGTGCGTACTCAGTGTGGAACACTGTATAGCCTTACGATCTGTGACT 796
Db 721 GAACGAGATGAACACTTGTGCGTACTCAGTGTGGAACACTGTATAGCCTTACGATCTGTGACT 780
OY 797 TATCCCATGTGTATTAACAAGTGTGTGACTGTGGCGGTGTATACGCTTCTTCTGACTTG 856
Db 781 TATCCCATGTGTATTAACAAGTGTGTGACTGTGGCGGTGTATACGCTTCTTCTGACTTG 840
OY 857 TCTAGTTGGGCGGAGTTTGTGAACCAAGCAAGGCTTACCTTGGCCATGAGCTGAGACT 916
Db 841 TCTAGTTGGGCGGAGTTTGTGAACCAAGCAAGGCTTACCTTGGCCATGAGCTGAGACT 900
OY 917 CGTTGTGCGGCTTTCACGTTCTGTGAGTCTTCTTCTATGTTGGCTGTGAAGGTTGGG 976
Db 901 CGTTGTGCGGCTTTCACGTTCTGTGAGTCTTCTTCTATGTTGGCTGTGAAGGTTGGG 954
OY 977 CCTCTCAAGGGCCCTGTGTGGGCTGTGAGGCAATGGCCAGAGGGGTATGAGCCAGACTGCT 1036
Db 955 CCTCTCAAGGGCCCTGTGTGGGCTGTGAGGCAATGGCCAGAGGGGTATGAGCCAGACTGCT 954
OY 1037 TGAGCAGAGGATGAGTGTGAGAAAGAGAGTCTCACGGGTAGAAAGAGCCAGGCGTG 1096
Db 955 TGAGCAGAGGATGAGTGTGAGAAAGAGAGTCTCACGGGTAGAAAGAGCCAGGCGTG 954
```

```
OY 1097 GTGGGCAACACCTGTATATCCAGTACTCGGAGGCTGAGGACAGAGAAATCGTTGAACC 1156
Db 955 GTGGGCAACACCTGTATATCCAGTACTCGGAGGCTGAGGACAGAGAAATCGTTGAACC 954
OY 1157 CGGAGGGGAGAGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1216
Db 955 CGGAGGGGAGAGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 997
OY 1217 GATTTTGAACAACAATGATGTGTGACAGGAAATTTGAGAGTGTCCCTGTGTGTGTGTGTGTGT 1276
Db 998 GATTTTGAACAACAATGATGTGTGACAGGAAATTTGAGAGTGTCCCTGTGTGTGTGTGTGTGT 1057
OY 1277 GAGATGCAACAGGACCTTCCCTCGGATGAGCCCGGACATGTACTGGAATAAGCCGAGCCA 1336
Db 1058 GAGATGCAACAGGACCTTCCCTCGGATGAGCCCGGACATGTACTGGAATAAGCCGAGCCA 1117
OY 1337 CAGCCCCCTTACACAGCTGTCTCCGCGCAATTCCTGTGAGGCTCTTTATGAGCTTCAAC 1396
Db 1118 CAGCCCCCTTACACAGCTGTCTCCGCGCAATTCCTGTGAGGCTCTTTATGAGCTTCAAC 1177
OY 1397 TTCAACATCAGCTGTGAACAAAGAGATGAGATTCCAGCCCAATCAGGAGGACAGAGAG 1456
Db 1178 TTCAACATCAGCTGTGAACAAAGAGATGAGATTCCAGCCCAATCAGGAGGACAGAGAG 1237
OY 1457 GATGCTCAGGCTGTGCAATTTGGCCGCTTCTTATGAGCTGTGAGTCCATATATCAATCTT 1516
Db 1238 GATGCTCAGGCTGTGCAATTTGGCCGCTTCTTATGAGCTGTGAGTCCATATATCAATCTT 1297
OY 1517 CCGAGGCAAACTCAAGAGCAAACTATGTGGCCCAAGAGGAAATCCCTTCTCAACAG 1576
Db 1298 CCGAGGCAAACTCAAGAGCAAACTATGTGGCCCAAGAGGAAATCCCTTCTCAACAG 1357
OY 1577 GGCCTGCCCAAAAACCAACAAGGACGCAACAAAGCTTTAGGGCCAGAGAAACAACAG 1636
Db 1358 GGCCTGCCCAAAAACCAACAAGGACGCAACAAAGCTTTAGGGCCAGAGAAACAACAG 1417
OY 1637 GCCTGAACTTAAAGCTGTGAGCGCTTCAAGTGTGGCCCACTGTATCAAGAGGCAAGG 1696
Db 1418 GCCTGAACTTAAAGCTGTGAGCGCTTCAAGTGTGGCCCACTGTATCAAGAGGCAAGG 1477
OY 1697 TACTACAGTGTCCCAACAAGGCGCTTCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1756
Db 1478 TACTACAGTGTCCCAACAAGGCGCTTCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1537
OY 1757 TCAGGCGGCTCAAACTTCAACAGTGTCAAGGCAATGACACCAAGAAAGCTTAAAG 1816
Db 1538 TCAGGCGGCTCAAACTTCAACAGTGTCAAGGCAATGACACCAAGAAAGCTTAAAG 1597
OY 1817 ACTGTGAGTTCTGGGCGCAAGAAAGTTTGAATTTGCTCTCAGAGAGCGATGGGCGCTTG 1876
Db 1598 ACTGTGAGTTCTGGGCGCAAGAAAGTTTGAATTTGCTCTCAGAGAGCGATGGGCGCTTG 1657
OY 1877 ATGAGGCAACCAAGAAATCTCAAGTGTGAGGAGGAAACTGTGAGTTTAACTTGAACGAGT 1936
Db 1658 ATGAGGCAACCAAGAAATCTCAAGTGTGAGGAGGAAACTGTGAGTTTAACTTGAACGAGT 1717
OY 1937 ATGCAAGATGCCCGGAAATCACTCAAGAACTTTGGAACAATCAACAACAACAATA 1996
Db 1718 ATGCAAGATGCCCGGAAATCACTCAAGAACTTTGGAACAATCAACAACAACAATA 1777
OY 1997 CACACTACCTCAAAAGATCACTGATGATCTTATTTGGGCGCTTGAAGAAACAGGAGTGAAGCA 2056
Db 1778 CACACTACCTCAAAAGATCACTGATGATCTTATTTGGGCGCTTGAAGAAACAGGAGTGAAGCA 1837
OY 2057 CATTCCTAACTGTCTTCTATGAGGAGATGTTTGGCCAGCAGGCTCTTCACTGTGTGTAC 2116
Db 1838 CATTCCTAACTGTCTTCTATGAGGAGATGTTTGGCCAGCAGGCTCTTCACTGTGTGTAC 1897
OY 2117 ACCAGCAGGACACTGTATCAGTCAACAGCCTTACAGCTGTGTCACTGAGAAAGCTGTCT 2176
Db 1898 ACCAGCAGGACACTGTATCAGTCAACAGCCTTACAGCTGTGTCACTGAGAAAGCTGTCT 1957
OY 2177 ACAACAGCTGATCAAAATGTATAGCTTAATGAATAATCCCAAGCTTACGCTT 2236
```

Db 1958 ACAACAGCCTGAATCAATAGTAGCTTAATAGATAAAATCCAGACTACTGAGCCTT 2017
Qy 2237 TAAAGCCCTTTATTCATATAAAACCTGTGAACCTAGACTGAACCATTTGAAACATTAACT 2296
Db 2018 TAAAGCCCTTTATTCATATAAAACCTGTGAACCTAGACTGAACCATTTGAAACATTAACT 2077
Qy 2297 CAGACTGTGATTCAGATTCGGGAACCTTAGTTCTATCTGAAATCAGACAGCCACACC 2356
Db 2078 CAGACTGTGATTCAGATTCGGGAACCTTAGTTCTATCTGAAATCAGACAGCCACACC 2137
Qy 2357 TTAGTATATCTGCCCAACTAATGAGTTTAATAAA 2390
Db 2138 TTAGTATATCTGCCCAACTAATGAGTTTAATAAA 2171

RESULT 6
CQ729633 2170 bp DNA linear PAT 03-FEB-2004
LOCUS Sequence 15567 from Patent WO02068579.
DEFINITION CQ729633
ACCESSION CQ729633
VERSION CQ729633.1 GI:42301252
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
TITLE Kites, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A, 15567 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES
source
1. .2170
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 80.0%; Score 1942.8; DB 6; Length 2170;
Best Local Similarity 91.3%; Pred. No. 0;
Matches 2168; Conservative 0; Mismatches 2; Indels 204; Gaps 2;

Qy 17 CTTAGTGGCCAGACCTTGTGGGATCATGGAACCCACTTGACCCAAAG 76
Db 1 CCTAGTGGCCAGACCTTGTGGGATCATGGAACCCACTTGACCCAAAG 60
Qy 77 CCCACCTGCTGACGCCACCTGCGCCATGACCATCACTTACACAGCCAAAGTGGCTAA 136
Db 61 CCCACCTGCTGACGCCACCTGCGCCATGACCATCACTTACACAGCCAAAG-GGCTAA 119
Qy 137 TGCCCGCTTAGAGCTCTTCTCCGCGCTGCTGCTGCTGCGGGGACATCTCAAGCT 196
Db 120 TGCCCGCTTAGAGCTCTTCTCCGCGCTGCTGCTGCTGCGGGGACATCTCAAGCT 179
Qy 197 GCTAATATGCGAGTTCTTAATCTTCTGCTGCTGCTACTACATCATCCGCTTATTTATAG 256
Db 180 GCTAATATGCGAGTTCTTAATCTTCTGCTGCTGCTACTACATCATCCGCTTATTTATAG 239
Qy 257 GCTGAGCCCTCAACGGAAGAAACAAGCTGATGTTTGAAGAACTGACTCTGTTATTCGACAG 316
Db 240 GCTGAGCCCTCAACGGAAGAAACAAGCTGATGTTTGAAGAACTGACTCTGTTATTCGACAG 299
Qy 317 CTACATCAGCTCATCCCATTTCTTCTGCTGCTGCGGCTTCTACGTTGACGCTGCTGAC 376
Db 300 CTACATCAGCTCATCCCATTTCTTCTGCTGCTGCGGCTTCTACGTTGACGCTGCTGAC 359
Qy 377 CCGCTGCTGGAACAAGTACGAAGAACTGCGCTGCGCGGACCGCTCATGAGCCTTGCTGTC 436
Db 360 CCGCTGCTGGAACAAGTACGAAGAACTGCGCTGCGCGGACCGCTCATGAGCCTTGCTGTC 419

Qy 437 GGGCTTGTTCGAGGCAAGGACGAGACAGGCGGCTGTGCGGCGACGCTATCCGCTA 496
Db 420 GGGCTTGTTCGAGGCAAGGACGAGACAGGCGGCTGTGCGGCGACGCTATCCGCTA 479
Qy 497 CGCCAACTGGGCAACGCTGCTCATCTGCGCAGCGTCAAGCACCGCAGTCTACAGCGCTT 556
Db 480 CGCCAACTGGGCAACGCTGCTCATCTGCGCAGCGTCAAGCACCGCAGTCTACAGCGCTT 539
Qy 557 CCCGAGGCGCCAGACCTGCTGCAAGCAGGCTTTATAGCTCCGCGAAGAACAGAGCTT 616
Db 540 CCCGAGGCGCCAGACCTGCTGCAAGCAGGCTTTATAGCTCCGCGAAGAACAGAGCTT 599
Qy 617 GGAAGAACTGAGCCTACACACAACATGTTCTGGGCGCTCGGGTGTGTTGGCAACCT 676
Db 600 GGAAGAACTGAGCCTACACACAACATGTTCTGGGCGCTCGGGTGTGTTGGCAACCT 659
Qy 677 GTCAATGAAAGGCTGTGGCTTGGAGGTGCAATCCGGAACCTATCTGCTCCAGAGCTGCT 736
Db 660 GTCAATGAAAGGCTGTGGCTTGGAGGTGCAATCCGGAACCTATCTGCTCCAGAGCTGCT 719
Qy 737 GAACGAGATGAACAACCTTGGCTACTCACTGCTGAGACACCTGATAGCTTACGATTTAG 796
Db 720 GAACGAGATGAACAACCTTGGCTACTCACTGCTGAGACACCTGATAGCTTACGATTTAG 779
Qy 797 TATCCACTGCTGTATACACAGGTGTGACCTGTGGCGGTGACAGCTTCTTCTGACTTG 856
Db 780 TATCCACTGCTGTATACACAGGTGTGACCTGTGGCGGTGACAGCTTCTTCTGACTTG 839
Qy 857 TCTAGTTGGGCGGAGTTTCTGAACCCAGCCAGGCGCTTACCTGCGCATGAGCTGACCT 916
Db 840 TCTAGTTGGGCGGAGTTTCTGAACCCAGCCAGGCGCTTACCTGCGCATGAGCTGACCT 899
Qy 917 CGTTGTGCGCGCTCTTCAAGCTTCTGCAAGTTCTTCTATATTTGGCTGGAAGTGG 976
Db 900 CGTTGTGCGCGCTCTTCAAGCTTCTGCAAGTTCTTCTATATTTGGCTGGAAGTGG 953
Qy 977 CCTCTCAGGCGCTGTGCTGAGGCTGGAAGCATGGCCAGAGGGGTATGACCGACAGCTGCT 1036
Db 954 ----- 953
Qy 1037 TGAGACGAGATGCGAGTGTACGAAAGGAAGTCTCACGGGTAGAAAGCAGCCAGCGTG 1096
Db 954 ----- 953
Qy 1097 GTGGCGCACCTGTATTCAGCTACTCGGGAAGGCTGAGGACGAAATGCTTGAACC 1156
Db 954 ----- 953
Qy 1157 CGGAGGCGGAGGTTGTGTGGGACAGAGCTCATCAACCCCTTTGAGAGAGATGATGAT 1216
Db 954 -----GTTGGCAGAGAGCTCATCAACCCCTTTGAGAGAGATGATGAT 996
Qy 1217 GATTTTGAACCACTGATTTGTGACAGAGAAATTTGACAGGATGTCCTGTTGGCTGAT 1276
Db 997 GATTTTGAACCACTGATTTGTGACAGAGAAATTTGACAGGATGTCCTGTTGGCTGAT 1056
Qy 1277 GAGATGACACGAGACCTGCTCGGATGAGACCGGACATGTAAGTGAATTAAGCCGAGCCA 1336
Db 1057 GAGATGACACGAGACCTGCTCGGATGAGACCGGACATGTAAGTGAATTAAGCCGAGCCA 1116
Qy 1337 CAGCCCGCTACACAGCTGCTCGGCCAGTTCGCTGAGAGCTCTTTATATGAGCTCCACC 1396
Db 1117 CAGCCCGCTACACAGCTGCTCGGCCAGTTCGCTGAGAGCTCTTTATATGAGCTCCACC 1176
Qy 1397 TTCAACATCAGCCTGAAACAAAGAGATGAGATTTCAGCCCAATCAGAGAGAGAGAG 1456
Db 1177 TTCAACATCAGCCTGAAACAAAGAGATGAGATTTCAGCCCAATCAGAGAGAGAGAG 1236
Qy 1457 GATGCTCAGCGTGCATCATTTGCGCTTCTTAGGCTGCAAGTCCATGATCACAATCTCT 1516
Db 1237 GATGCTCAGCGTGCATCATTTGCGCTTCTTAGGCTGCAAGTCCATGATCACAATCTCT 1236
Qy 1517 CCCAGGCGAAACTGAAGACCAAACTACTGTGCGCCAAAGGGAATCCCTTCTCAAGAG 1576

|||||
Db 1297 CCGAGGCGCAACCTCAAGGACCAAACTACTGTGGCCCAAGGGAGATCCCTTCTCCACGAG 1356
Oy 1577 GGCGTGCCTCAAAAACCAAGGCGAGCCCAACAGAACTTTAGGGCCGAGGAAGCAACAG 1636
Db 1357 GGCGTGCCTCAAAAACCAAGGCGAGCCCAACAGAACTTTAGGGCCGAGGAAGCAACAG 1416
Oy 1637 GGCGTGCCTCAAAAACCAAGGCGAGCCCTTCAAGCTGTGGCCCACTGTATCGAGGGCCAGG 1696
Db 1417 GGCGTGCCTCAAAAACCAAGGCGAGCCCTTCAAGCTGTGGCCCACTGTATCGAGGGCCAGG 1476
Oy 1697 TACTACAGTGGCCCAAGAGCCCTTCAAGCCCACTTCCATGTTCTTCCCTTGAAGCA 1756
Db 1477 TACTACAGTGGCCCAAGAGCCCTTCAAGCCCACTTCCATGTTCTTCCCTTGAAGCA 1536
Oy 1757 TCAGCCGCGTCAAAAGCTTCAAGTGTCAAGGATAGACCAACCAAAAGCTTAAAG 1816
Db 1537 TCAGCCGCGTCAAAAGCTTCAAGTGTCAAGGATAGACCAACCAAAAGCTTAAAG 1596
Oy 1817 ACTGTAGATTCTGGGGCCAAAGAAAGTTTGAATTGCTCTCAAGAGAGAGTGGGGCTTG 1876
Db 1597 ACTGTAGATTCTGGGGCCAAAGAAAGTTTGAATTGCTCTCAAGAGAGAGTGGGGCTTG 1656
Oy 1877 ATGAGACACCCAGAAATCTCAAGTGAAGAGAAAGCTGTGAGTTTAACTTGA 1936
Db 1657 ATGAGACACCCAGAAATCTCAAGTGAAGAGAAAGCTGTGAGTTTAACTTGA 1716
Oy 1937 ATGCGAGAGATCCCGCAAAATCACTCAAGAAACCTTTGAAACATCAACCAACATA 1996
Db 1717 ATGCGAGAGATCCCGCAAAATCACTCAAGAAACCTTTGAAACATCAACCAACATA 1776
Oy 1997 CACACACACCTCAAAAGATCACTGATCTTTATTTGGGCTTTGAAAGAGGATGAACA 2056
Db 1777 CACACACACCTCAAAAGATCACTGATCTTTATTTGGGCTTTGAAAGAGGATGAACA 1836
Oy 2057 CATTCCTAAGCTGTCTCTAATGGAGAGTCTTGGCAGCGAGTCTCACTGTGTAC 2116
Db 1837 CATTCCTAAGCTGTCTCTAATGGAGAGTCTTGGCAGCGAGTCTCACTGTGTAC 1896
Oy 2117 ACCAGAGAGACCTGATTCAGTCAAGCCATACAGCTGTCCACACTGAAGAGCTGTCT 2176
Db 1897 ACCAGAGAGACCTGATTCAGTCAAGCCATACAGCTGTCCACACTGAAGAGCTGTCT 1956
Oy 2177 ACAACAGCTGATCAAAATGTTAGCTTAATAGTAAATAATCCAGACTTCTCAAGCTT 2236
Db 1957 ACAACAGCTGATCAAAATGTTAGCTTAATAGTAAATAATCCAGACTTCTCAAGCTT 2016
Oy 2237 TAATGCTTTTATTCATTAATAAATCTGTGAAGCTAGACTGAACATTTGAAT 2296
Db 2017 TAATGCTTTTATTCATTAATAAATCTGTGAAGCTAGACTGAACATTTGAAT 2076
Oy 2297 CAGACTCTGAGATTCAAGATCGGGAACCTTTAGTTCTATCTGAATCCAAAGCAGCACCC 2356
Db 2077 CAGACTCTGAGATTCAAGATCGGGAACCTTTAGTTCTATCTGAATCCAAAGCAGCACCC 2136
Oy 2357 TTAGTATACGCCCCAACTAATAGTTTAATAAA 2390
Db 2137 TTAGTATACGCCCCAACTAATAGTTTAATAAA 2170
RESULT 7
AY357925 2187 bp mRNA linear PRI 26-AUG-2003
LOCUS Macaca fascicularis beestrophin (VMD2) mRNA, complete cds.
DEFINITION
ACCESSION AY357925
VERSION AY357925.1 GI:34013782
KEYWORDS
SOURCE Macaca fascicularis (crab-eating macaque)
ORGANISM Macaca fascicularis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecinae; Macaca.
REFERENCE
1 (bases 1 to 2187)

AUTHORS Okamoto, H., Umeda, S., Suzuki, M.T., Yoshikawa, Y., Tanaka, Y. and Iwata, T.
TITLE Molecular Cloning of VMD2 Gene from Cynomolgus Monkey (Macaca fascicularis)
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2187)
AUTHORS Okamoto, H., Umeda, S., Suzuki, M.T., Yoshikawa, Y., Tanaka, Y. and Iwata, T.
TITLE Direct Submission
JOURNAL Submitted (01-AUG-2003) National Tokyo Medical Center, National Institute of Sensory Organs, 2-5-1, Higashi-choke, Meguro-ku, Tokyo 152-0021, Japan
FEATURES
source
location/Qualifiers
1..2187
/organism="Macaca fascicularis"
/mol_type="mRNA"
/db_xref="taxon:9541"
1..2187
/gene="VMD2"
67..1824
/gene="VMD2"
/product="beestrophin"
/codon_start=1
/protein_id="AA056049.1"
/db_xref="GI:34013783"
/translation="MTITTSQVANRLGSFRLLICMRGSIYKLLYGEFFIFLLCY IIRFLRLVTEOQUMFERKLYICDSYIQLPISEVLYFVTLVTRMNDYENIPW PDRMLSVSGFVEGKDEQRLRLIRVANIENVILRSVSTAYAKFPPSAOHVOA GCMTPAEHDKLEKLSLPHNMFVWPMVPMFLSMKMLGGRIIDPILLOSILNEMKTL TFLOPFYVGMILKVAEOLINPFGEDDDDEETWVIDRINOVSLIADVMKHDTLPMEP DMYNEPEHPPTTAASQFRASFMGSPFNILANKEMEPFNOEDKDAATGILIG FIGLOSHHPRPGANSRTKLIPKXSLHLBELPKHKAIVKONVGLSDNRMKIKAV DAKSPILYORPQYISAFQTPUSPTPMFPPEPSVLKLSHTVGLIDTOKSLKTVSSG ANKSELSGSQALMEHPEVSHVRKVEFNLIDMPIIPENHLKEPLELSTYNIHAT LNDHDPYVALENRDEAHS"
ORIGIN
Query Match 74.5%; Score 1809.6; DB 9; Length 2187;
Best Local Similarity 87.5%; Pred. No. 0;
Matches 2092; Conservative 0; Mismatches 89; Indels 209; Gaps 2;
Oy 39 GGATCATCGGACCCCACTGGAACCCCACTGACCCCAACCCCACTGTGCAAGCCCACTGC 98
Db 1 GGATCATCGGATCCACCTGGAACCCCACTGACCCCAACCCCACTGTGCAAGAGTGC 60
Oy 99 CTGGCCATGACATCACTTAACAAGCCAAAGTGCTAAATGCCGCTTAAGCTCTCTCC 158
Db 61 CTGGCCATGACATCACTTAACAAGCCAAAGTGCTAAATGCCGCTTAAGCTCTCTCC 120
Oy 159 CGCTGTGCTGTGCTGCGGGGGCAGATCTACAAGCTGTATATGGAGTGTCTTAATC 218
Db 121 CGCTGTGCTGTGCTGCGGGGGCAGATCTACAAGCTGTATATGGAGTGTCTTAATC 180
Oy 219 TTCTGTCTGTCTACTATCATCATCGCTTTATTTATAGCTGTGCGCTTCAAGGAACAA 278
Db 181 TTCTGTCTGTCTACTATCATCATCGCTTTATTTATAGCTGTGCGCTTCAAGGAACAA 240
Oy 279 CAGCTGATGTTTGAAGAACTGATCTGTATTCGACAGCTACATCCAGCTATCCCAATT 338
Db 241 CAGCTGATGTTTGAAGAACTGATCTGTATTCGACAGCTACATCCAGCTATCCCAATT 300
Oy 339 TCCTTCGTGTGAGGCTTCTACGTGACGCTGTGCTGACCCGCTGTGTGAACCAATACAG 398
Db 301 TCCTTCGTGTGAGGCTTCTACGTGACGCTGTGCTGACCCGCTGTGTGAACCAATACAG 360
Oy 399 AACCTGCGTGGCCCAACCGCTTATGAGCTGTGTGCGGCTTTCGTGAAGCAAGAC 458
Db 361 AACCTGCGTGGCCCAACCGCTTATGAGCTGTGTGCGGCTTTCGTGAAGCAAGAC 420
Oy 459 GAGCAAGCGCGGCTGTGCGGCGCAGCGCTCATCGCTACGACCAACCTGGGCAACGCTGC 518
Db 421 GAGCAAGCGCGGCTGTGCGGCGCAGCGCTCATCGCTACGACCAACCTGGGCAACGCTGC 480

QY 519 ATCTGCGCAGCGTCAGCAGCCAGTCTACAGCGCTTCCCGAGGCCAGCAGCTGTG 578
DB 481 ATCTTGCGCAGCGTCAGCAGCCAGTCTACAGCGCTTCCCGAGGCCAGCAGCTGTG 540
QY 579 CAAGCAGGCTTTATGACTCCGCGAGAACACAGCAGTTGAGAAACTGAGCTTACAC 638
DB 541 CAACAGGCTTTATGACTCCGCGAGAACACAGCAGTTGAGAAACTGAGCTTACAC 600
QY 639 AACATGTTCTGGGTGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 698
DB 601 AACATGTTCTGGGTGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660
QY 699 GGTGGAATCCGGGAGCCCTATCTGTCTGAGAGCCGTGAGAGAGAGATGAACTGTGGT 758
DB 661 GGTGGAATCCGGGAGCCCTATCTGTCTGAGAGCCGTGAGAGAGATGAACTGTGGT 720
QY 759 ACTGAGTGTGACACTGTGTATGCTTACAGCTGTGATTAATCCACTGTGTATACAG 818
DB 721 ACTGAGTGTGACACTGTGTATGCTTACAGCTGTGATTAATCCACTGTGTATACAG 780
QY 819 GTGTGATCTGTGCGGTGTACAGCTTCTTCTGATCTTGTGTGTGTGTGTGTGTGTGT 878
DB 781 GTGTGATCTGTGCGGTGTACAGCTTCTTCTGATCTTGTGTGTGTGTGTGTGTGTGT 840
QY 879 AACCCAGCCAGAGCCCTTACCTGTGAGCTGAGCTGTGTGTGTGTGTGTGTGTGTGT 938
DB 841 AACCCAGCCAGAGCCCTTACCTGTGAGCTGAGCTGTGTGTGTGTGTGTGTGTGTGT 900
QY 939 CTGAGATTCTTCTTATGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 998
DB 901 CTGAGATTCTTCTTATGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 932
QY 999 TGGAGGATGTGCGCAGAGGGGTTCATGTGCGACAGCTGTGTGTGTGTGTGTGTGT 1058
DB 933 ----- 932
QY 1059 GAAAGGAAGTCTACCGGGTGAAGAACAGCAGCGGTGTGTGTGTGTGTGTGTGTGT 1118
DB 933 ----- 932
QY 1119 GCTACTCGGAGGCTGAGCAGAGAAATCGCTTGAACCCGGGAGCGGAGTTGTGTGT 1178
DB 933 ----- 932
QY 933 ----- 932
DB 933 ----- 932
QY 1179 CAGAGCAGCTATCAACCTTTTGGAGAGATGATGATTTTGAAGCAACTGTGATTTG 1238
DB 938 CAGAGCAGCTATCAACCTTTTGGAGAGAGAGATGATGATTTTGAAGCAACTGTGATTTG 997
QY 1239 TCGACAGGAATTTTGCAGGTTGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1298
DB 998 TCGACAGGAATTTTGCAGGTTGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1057
QY 1299 GGATGAGCCGAGACATGTACTGAAATTAAGCCGAGCCAGCCCTTACACAGCTGTGT 1358
DB 1058 GGATGAGCCGAGACATGTACTGAAATTAAGCCGAGCCAGCCCTTACACAGCTGTGT 1117
QY 1359 CCGCCCACTTCGTCGAGCTTCTTTATGGGCTTCACTTCAACATCACTGCTGAACAA 1418
DB 1118 CCGCCCACTTCGTCGAGCTTCTTTATGGGCTTCACTTCAACATCACTGCTGAACAA 1177
QY 1419 AGGAGATGGAATTTTGCAGCCCAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1478
DB 1178 AGGAGATGGAATTTTGCAGCCCAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1237
QY 1479 GCGGCTTCTAGAGCTGTGAGTCCCATGTATCAACATCTTCCAGAGGCAAACTCAAGACA 1538
DB 1238 GCGGCTTCTAGAGCTGTGAGTCCCATGTATCAACATCTTCCAGAGGCAAACTCAAGACA 1297
QY 1539 AACTACTGTGGCCCAAGAGGAGATCTTCTTCCAGAGGAGCTGTCCCAAAAACCAAG 1598
DB 1298 AACTACTGTGGCCCAAGAGGAGATCTTCTTCCAGAGGAGCTGTCCCAAAAACCAAG 1357

QY 1599 CAGCCAAACAGACGTTAGGGGCCAGAGAGAACACAGAGCCCTGAGAGCTTAAAGCTGTG 1658
DB 1358 CAGTCAAAACAGACGTTAGGGGCCAGAGAGAACACAGAGCCCTGAGAGCTTAAAGCTGTG 1417
QY 1659 ACGCTTCAAGTCTGGGCCACTGTATCAAGAGCCAGGCTTACTACGTGCCCCACAGAGC 1718
DB 1418 ATGCTTCAAGTCTGGGCCACTGTATCAAGAGCCAGGCTTACTACGTGCCCCACAGAGC 1477
QY 1719 CCCTCAGCCCACTCCAGTGTCTTCCCTTGAAGACCATCAGCCGCGTCAAAAGCTTCA 1778
DB 1478 CCCTCAGCCCACTCCAGTGTCTTCCCTTGAAGACCATCAGCTGTGTCAAAAGCTTCA 1537
QY 1779 GTGTCAAGAGATAGACACCAAGAGCAAAAGCTTAAAGACTGTGATTTCTGGGCCAAGA 1838
DB 1538 GTGTCAAGAGATAGACACCAAGAGCAAAAGCTTAAAGACTGTGATTTCTGGGCCAAGA 1597
QY 1839 AAAGTTGAATTTGCTCTCAGAGAGCGATGGGGCTTGTATGAGACCCAGAGATATCTC 1898
DB 1598 AAAGTTGAATTTGCTCTCAGAGAGCGATGGGGCTTGTATGAGACCCAGAGATATCTC 1657
QY 1899 AAGTGAAGAGAGAACTGTGAGTTTAACTGAGGATATGCGCAGAGATCCCGAATAATC 1958
DB 1658 AAGTGAAGAGAGAACTGTGAGTTTAACTGAGGATATGCGCAGAGATCCCGAATAATC 1717
QY 1959 AACTCAAAAGACCTTTGGAGCAATCAACCAACATACACTACACTCAAAAGATCA 2018
DB 1718 AACTCAAAAGACCTTTGGAGCAATCAACCAACATACACTACACTCAAAAGATCA 1777
QY 2019 TGGATCTTATTTGGGCTTGGAGAAACAGAGATGAAGACATTTCTTACCTGCTTCTTAT 2078
DB 1778 TGGATCTTATTTGGGCTTGGAGAAACAGAGATGAAGACATTTCTTACCTGCTTCTTAT 1837
QY 2079 GGGAGTGTTCGCCAGCCAGGCTTCTCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2138
DB 1838 GGGAGTGTTCGCCAGCCAGGCTTCTCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1897
QY 2139 CACAGCCATACAGCTGTCTCAGCTGAGAGAGTGTCTTACCAACAGCTGAAATCAATGAT 2198
DB 1898 CACAGCCATACAGCTGTCTCAGCTGAGAGAGTGTCTTACCAACAGCTGAAATCAATGAT 1957
QY 2199 TAGCTTAATAGATTA-----AATCCAGACTACTTCAAGCTTTAATGCTTTTATTTCA 2252
DB 1958 TAGCTTAATAGATTA-----AATCCAGACTACTTCAAGCTTTAATGCTTTTATTTCA 2017
QY 2253 TAAAGACGTGAAGCTTGAAGTGAAGCAATTTGAAGCAATTTGAAGCTTGAAGCTTGAAG 2312
DB 2018 TAAAGACGTGAAGCTTGAAGTGAAGCAATTTGAAGCAATTTGAAGCTTGAAGCTTGAAG 2077
QY 2313 AGTGGGAAACCTTATGTTCTATCTGATCTGAATCAAGACAGCCACTTATATCTGCCAA 2372
DB 2078 AGTGGGAAACCTTATGTTCTATCTGATCTGAATCAAGACAGCCACTTATATCTGCCAA 2137
QY 2373 ACTAATGATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2422
DB 2138 ACTAATGATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2187

RESULT 8
BC015220 2441 bp mRNA linear PRI 07-OCT-2003
LOCUS BC015220.2 GI:34783515
DEFINITION Homo sapiens vitelliform macular dystrophy (Best disease, bestrophin), mRNA (cdna clone MGC:117761 IMAGE:3877806), complete cds.
ACCESSION BC015220
VERSION BC015220.2
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 2441)
Straubeberg,R.L., Feingold,E.A., Grouse,L.H., Dergs,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenneman,C.M., Schuler,G.D.,

Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stedman, M., Soares, M.B., Donald, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Udell, T.B., Toshiyuki, S., Caminici, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., Wierley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Sanchez, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smalls, D.E., Schermer, A., Schein, J.E., Jones, S.J., and Marra, M.A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

22388257

JOURNAL MEDLINE PubMed 12477932

2 (bases 1 to 2441)

Strausberg, R.

Direct Submission

Submitted (01-OCT-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

On Sep 16, 2003 this sequence version replaced gi:21955361.

Contact: MGC help desk

Email: gcgabs-remail.nih.gov

Tissue Procurement: DCTD/DTF/Gazdar

cDNA Library Preparation: Life Technologies, Inc.

DNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)

DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www.shgc.stanford.edu>

Contact: (Dickson, Mark) mcd@paxil.stanford.edu

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LIML at: <http://image.llnl.gov>

Series: IRAX Plate: 14 Row: B Column: 9.

location/Qualifiers

1. 2441

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:17761 IMAGE:3877806"

/issue_type="lung, large cell carcinoma"

/clone_lib="NIH_MGC_68"

/lab_host="DH10B"

/note="vector: pCMV-SPORT6"

1. 2441

/gene="VMD2"

/note="synonyms: BMD, BERT"

/db_xref="locusid:7439"

/db_xref="MIM:153700"

573. 2069

/gene="VMD2"

/codon_start=1

/product="VMD2 protein"

/protein_id="AAH15220.1"

/db_xref="GI:22800661"

/db_xref="LocusID:7439"

/db_xref="MIM:153700"

/translation="MEKLTLYCDSYTIPLISFVIGFYVTVIWMNOYENLPMPPDLMSIVSGFEVGEDEGRILRLIRVANSNTLIRSVSTAYRFRPAQHVDVGFMPBAHKOLEKSLPHNMFVPPVWPAISNTKAMUGKRLDPLLOSLLNENNTLTQCGHLYADWISLIVTVQVTVAVSFPLTCLVGRQFLNPAVYPHEHIDLVPVPTTLOFFFYVGMKLVSLLAVDEMHODLPMEEDPMWNKDEPPPYTAASACFRASFMGST

ORIGIN

Query Match 65.3%; Score 1587; DB 9; Length 2441;

Beet Local Similarity 86.7%; Pred. No. 0;

Matches 1884; Conservative 0; Mismatches 5; Indels 284; Gaps 1;

misc_feature

576. 1427

/gene="VMD2"

/note="DUF289, Region: Putative membrane protein. This family called family 8 in, may be a transmembrane protein The specific function of this protein is unknown"

/db_xref="CDD:pfam01062"

257 GCTGACCTCAGGAAAGCAAGAGCAAGAGCGCTGCTGAGCGAGCGCTATGAGCTGCGAG 316

545 GCTGGCCCTACGGAAGAACAGAGCTGATGTTGAGAACTGATCTGATTTGCGACAG 604

317 CTACATCAGCTCATCCCATTTCTTCTGCTGAGGCTTCTAGCTGACGCTGCTGAC 376

605 CTACATCAGCTCATCCCATTTCTTCTGCTGAGGCTTCTAGCTGACGCTGCTGAC 664

377 CCGCTGTGAGAACAGAGAACAGAGAACAGAGCGCTGCTGAGCGCTGCTGAGCTGCTG 436

665 CCGCTGTGAGAACAGAGAACAGAGAACAGAGCGCTGCTGAGCGCTGCTGAGCTGCTG 724

437 GAGCTGTGAGAACAGAGAACAGAGAACAGAGCGCTGCTGAGCGCTGCTGAGCTGCTG 496

725 GAGCTGTGAGAACAGAGAACAGAGAACAGAGCGCTGCTGAGCGCTGCTGAGCTGCTG 784

497 GCGCAACCTGAGAACAGAGAACAGAGAACAGAGCGCTGCTGAGCGCTGCTGAGCTGCTG 556

785 GCGCAACCTGAGAACAGAGAACAGAGAACAGAGCGCTGCTGAGCGCTGCTGAGCTGCTG 844

557 CCGCAACCTGAGAACAGAGAACAGAGAACAGAGCGCTGCTGAGCGCTGCTGAGCTGCTG 616

845 CCGCAACCTGAGAACAGAGAACAGAGAACAGAGCGCTGCTGAGCGCTGCTGAGCTGCTG 904

617 GAGAAACCTGAGAACAGAGAACAGAGAACAGAGCGCTGCTGAGCGCTGCTGAGCTGCTG 676

905 GAGAAACCTGAGAACAGAGAACAGAGAACAGAGCGCTGCTGAGCGCTGCTGAGCTGCTG 964

677 GTCATGAGAGCGCTGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 736

965 GTCATGAGAGCGCTGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 1024

737 GAAAGATGAAACCTGCTGAGTCACTGAGTGAACCTGATGCTGATGCTGATGCTGATGCTG 796

1025 GAAAGATGAAACCTGCTGAGTCACTGAGTGAACCTGATGCTGATGCTGATGCTGATGCTG 1084

797 TATCCACAGTGTATATACAGAGTGTATATATATATATATATATATATATATATATATAT 856

1085 TATCCACAGTGTATATACAGAGTGTATATATATATATATATATATATATATATATATAT 1144

857 TCTAGTGTGCGCGAGTGTGAAACCGAGAGCGCTGAGCGCTGAGCGCTGAGCGCTGAGCGCT 916

1145 TCTAGTGTGCGCGAGTGTGAAACCGAGAGCGCTGAGCGCTGAGCGCTGAGCGCTGAGCGCT 1204

917 CGTGTGTGCGCGCTGCTGAGCGCTGAGCGCTGAGCGCTGAGCGCTGAGCGCTGAGCGCTG 976

1205 CGTGTGTGCGCGCTGCTGAGCGCTGAGCGCTGAGCGCTGAGCGCTGAGCGCTGAGCGCTG 1257

977 CCTTCCAGGGCGCTGCTGAGCGCTGAGCGCTGAGCGCTGAGCGCTGAGCGCTGAGCGCTG 1036

1258 ----- 1257

1037 TGAAGCAGAGTCACTGCTCAGAGAGAGAGTCTCAGCGTAGAAGAGAGCGAGCGCTG 1096

1258 ----- 1257

1097 GTGGCGGACACCTGATCCAGCTACTCGGAGAGCTGAGAGAGAGAGTCTGAGAGCTGAGAGCT 1156

Db 1258 ----- 1257
 Qy 1157 CGGAGGCGGAGTTGTGTGGCAGAGAGCTCATCAACCCCTTTGGAGAGATGATGAT 1216
 Db 1258 ----- 1257
 Qy 1217 GATTTTGAGACCAATGAGATTGTGCAGAGAAATTTGACGTGTCCCTGTGTGGCTGTGAT 1216
 Db 1258 ----- 1280
 Qy 1277 GAGATGACACGAGACTGCTCTCGAGTGGAGCCCGACATGTACTGTGAATPAAGCCGAGCCA 1336
 Db 1281 GAGATGACACGAGACTGCTCTCGAGTGGAGCCCGACATGTACTGTGAATPAAGCCGAGCCA 1340
 Qy 1337 CAGCCCCCTTACAGAGCTGTCCGCGCAGTTCCGTGAGCTCCTCTTAATGGGCTCCACC 1336
 Db 1341 CAGCCCCCTTACAGAGCTGTCCGCGCAGTTCCGTGAGCTCCTCTTAATGGGCTCCACC 1400
 Qy 1397 TTCAACATCAGCTGTAACTCAAGAGAGATGAGTTCCAGCCCAATCAGAGAGAGAG 1456
 Db 1401 TTCAACATCAGCTGTAACTCAAGAGAGATGAGTTCCAGCCCAATCAGAGAGAGAG 1460
 Qy 1457 GATGCTCAGCTGTGATGAGTGGCGGCTTCTTGAAGCTGAGTCCCATGATCAGATCCT 1516
 Db 1461 GATGCTCAGCTGTGATGAGTGGCGGCTTCTTGAAGCTGAGTCCCATGATCAGATCCT 1520
 Qy 1517 CCCAGGCGCAACTCAAGAGACCAAACTACTGTGGCCCAAGAGAGAACTCCTTCTCAGCAG 1576
 Db 1521 CCCAGGCGCAACTCAAGAGACCAAACTACTGTGGCCCAAGAGAGAACTCCTTCTCAGCAG 1580
 Qy 1577 GGCCTGCGCCAAACCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1636
 Db 1581 GGCCTGCGCCAAACCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1640
 Qy 1637 GGCCTGAGAGCTTAAAGCTGTGAGAGCGCTTCAAGTCTGAGCCCACTGTATCAGAGGCCAGC 1696
 Db 1641 GGCCTGAGAGCTTAAAGCTGTGAGAGCGCTTCAAGTCTGAGCCCACTGTATCAGAGGCCAGC 1700
 Qy 1697 TACTACAGTGGCCCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1756
 Db 1701 TACTACAGTGGCCCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1760
 Qy 1757 TCAGAGCGGTCAAAAGCTTCAAGTGTCAAGGATGACAGCAACAAAGAGAGAGAGAG 1816
 Db 1761 TCAGAGCGGTCAAAAGCTTCAAGTGTCAAGGATGACAGCAACAAAGAGAGAGAGAG 1820
 Qy 1817 ACTGTGAGTCTGTGGGCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1876
 Db 1821 ACTGTGAGTCTGTGGGCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1880
 Qy 1877 ATGAGAGACCCAGAGAGATGTCTCAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 1936
 Db 1881 ATGAGAGACCCAGAGAGATGTCTCAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 1940
 Qy 1937 ATGCGAGAGATCCCGAGAGAGATCACTCAAGAGAGAGAGAGAGAGAGAGAGAGAG 1996
 Db 1941 ATGCGAGAGATCCCGAGAGAGATCACTCAAGAGAGAGAGAGAGAGAGAGAGAGAG 2000
 Qy 1997 CACACTCACTCAAAAGATCAGATGATCTTATTGGGCTTGGAGAGAGAGAGAGAGAG 2056
 Db 2001 CACACTCACTCAAAAGATCAGATGATCTTATTGGGCTTGGAGAGAGAGAGAGAGAG 2060
 Qy 2057 CATTCTCACTGCTCTTCTTAATGGGAGATCTTCCGACGCGAGTCTCACTGTGTGTAC 2116
 Db 2061 CATTCTCACTGCTCTTCTTAATGGGAGATCTTCCGACGCGAGTCTCACTGTGTGTAC 2120
 Qy 2117 ACCAGAGAGAGAGAGAGATGATCAAGTCAAGCAGATGATGATGATGATGATGATGAT 2176
 Db 2121 ACCAGAGAGAGAGAGATGATCAAGTCAAGCAGATGATGATGATGATGATGATGATGAT 2180
 Qy 2177 ACAACAGCTGATCAAAATGAGTTAGTTAATGATTAATAATCCAGACTACTGAGCCTT 2236

Db 2181 ACAACAGCTGATCAAAATGAGTTAGCTTAATGATTAATAATCCAGACTACTGAGCCTT 2240
 Qy 2237 TAATGCCCTTTTATTCATTAATAAATCTGTGAAGAGTGAAGTGAAGAGAGAGAGAG 2296
 Db 2241 TAATGCCCTTTTATTCATTAATAAATCTGTGAAGAGTGAAGTGAAGAGAGAGAGAG 2300
 Qy 2297 CAGACTGTGATGAGAGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2356
 Db 2301 CAGACTGTGATGAGAGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2360
 Qy 2357 TTAGTATTAATGAG 2416
 Db 2361 TTAGTATTAATGAG 2420
 Qy 2417 AAAAAAAAAAAAAA 2429
 Db 2421 AAAAAAAAAAAAAA 2433

RESULT 9
 LOCUS AY515704 1758 bp mRNA linear PRI 28-JAN-2004
 DEFINITION Homo sapiens beetrophin 1 mRNA, complete cds.
 ACCESSION AY515704
 VERSION AY515704.1 GI:41216872
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1758)
 AUTHORS Sun,H., Tsunenari,T., Yau,K.W., and Nathans,J.
 TITLE The vitelliform macular dystrophy protein defines a new family of chloride channels
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (6), 4008-4013 (2002)
 MEDLINE 21902086
 PUBMED 11904445
 REFERENCE 2 (bases 1 to 1758)
 AUTHORS Tsunenari,T., Sun,H., Williams,J., Cahill,H., Smallwood,P., Yau,K.W. and Nathans,J.
 TITLE Structure-function analysis of the beetrophin family of anion channels
 JOURNAL J. Biol. Chem. 278 (42), 41114-41125 (2003)
 MEDLINE 22917481
 PUBMED 12907679
 REFERENCE 3 (bases 1 to 1758)
 AUTHORS Sun,H., Tsunenari,T., Yau,K.-W. and Nathans,J.
 TITLE Direct Submission
 JOURNAL Submitted (02-JAN-2004) Molecular Biology, Neuroscience, Johns Hopkins University, 725 N. Wolfe Street, Baltimore, MD 21205, USA
 FEATURES
 source
 1..1758
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 1..1758
 /note="chloride channel"
 /codon_start=1
 /product="beetrophin 1"
 /protein_id="AA89654.1"
 /db_xref="GI:41216873"
 /translation="MTITYSOVNANRLGSPSRLLCLMRGSIYKLYLGEFLLPILCY IIRIVLALTEBQMLPEKTLTLCDSYICLIPSPFGPVPVTLVTRPMNOYENLPM PDRMLSVSGFVBGKDEQRLRRLTLIRYALGNVLLRVSIVTVYKRFPSAQHLYOA GFMTPAEHKOLEKLSLPHNMFVWVFWFANLSMKAWLIGRLDILLOSILNENKTLR TOCHILAYDMKISLPLVYQVTVAVVSFELTCLVGFOPLPAPAYAGHEDLVVPPV TPLQFPFYVGMILKVAEOLINPFGEDDDPFENMLVDNLLQVSLAVBEMHODLRMBP DMWNKREPPRPYTAASAPRRASFMGSEFNISLNKEMEPQPOBDEBDNAHIGIR FLGLOSHDHRPPRANSTKLMPRESLIRBGLPKHKAAKONVRGQEDNKAKYLVK DAFKSAPLYRPGYISAPQTPLESTPMFPLBPSAPSKHLSVYQIDTKDSLYKTVSSG AKKSFELSSDGLALMEHPESQVRRTVEFNLTDMPELIPENHLKEPLEOSPTNIHTT LKDHMDPYMALENRDEAHS"
 ORIGIN

Query Match 63.5%; Score 1543.4; DB 9; Length 1758;
Best Local Similarity 89.6%; Pred. No. 0;
Matches 1757; Conservative 0; Mismatches 1; Indels 203; Gaps 1;

```
QY 105 ATGACATCACTTACACAAAGCCAAAGGCTTAATCCCGCTTAAGCTCTTCTCCGCGCTG 164
Db 1 ATGACATCACTTACACAAAGCCAAAGGCTTAATCCCGCTTAAGCTCTTCTCCGCGCTG 60
QY 165 CTGCTGCTGCTGCGGGGAGAGATCTACAGCTGTATATGCGAGTCTTAAATCTTCCTG 224
Db 61 CTGCTGCTGCTGCGGGGAGAGATCTACAGCTGTATATGCGAGTCTTAAATCTTCCTG 120
QY 225 CTGCTGCTGCTGCGGGGAGAGATCTACAGCTGTATATGCGAGTCTTAAATCTTCCTG 284
Db 121 CTGCTGCTGCTGCGGGGAGAGATCTACAGCTGTATATGCGAGTCTTAAATCTTCCTG 180
QY 285 ATGTTTGAAGAACTGATCTGTATTTGAGCAAGCTTACATCCCAATTTCTTC 344
Db 181 ATGTTTGAAGAACTGATCTGTATTTGAGCAAGCTTACATCCCAATTTCTTC 240
QY 345 GTGCTGCGCTTCTACGTGAGCTGTGTCGTCACCGGCTTGAGAACAGTACAGAACTTG 404
Db 241 GTGCTGCGCTTCTACGTGAGCTGTGTCGTCACCGGCTTGAGAACAGTACAGAACTTG 300
QY 405 CCGTGGCCGACCGCTCATGAGCTGTGTGCGAGCTTGTGCAAGGCAAGGACGAGCAA 464
Db 301 CCGTGGCCGACCGCTCATGAGCTGTGTGCGAGCTTGTGCAAGGCAAGGACGAGCAA 360
QY 465 GCGCGCTGCTGCGGCGACGCTCATCCGCTACCGCAACCTGGGCAAGCTGTCTACTCTG 524
Db 361 GCGCGCTGCTGCGGCGACGCTCATCCGCTACCGCAACCTGGGCAAGCTGTCTACTCTG 420
QY 525 CCGAGCGTGAACAACGCACTTACCAAGCGCTTCCCAAGCGCCACCACTGGTGAAGCA 584
Db 421 CCGAGCGTGAACAACGCACTTACCAAGCGCTTCCCAAGCGCCACCACTGGTGAAGCA 480
QY 585 GGGTTATATGATCCCGGCAAGCAAGCAAGCTTGGAGAACTGAGCTTACCAACAACATG 644
Db 481 GGGTTATATGATCCCGGCAAGCAAGCAAGCTTGGAGAACTGAGCTTACCAACAACATG 540
QY 645 TTCTGGGTGCTGCGGTGTGTTTCCCAACTGTCAATGAAGCGTGGAGTGAAGTGA 704
Db 541 TTCTGGGTGCTGCGGTGTGTTTCCCAACTGTCAATGAAGCGTGGAGTGAAGTGA 600
QY 705 ATCCGGAACCTTATCTCTGCTCCAGGCTGCTGAACGAGATGAACACCTTGCATCTAG 764
Db 601 ATCCGGAACCTTATCTCTGCTCCAGGCTGCTGAACGAGATGAACACCTTGCATCTAG 660
QY 765 TGTGGAACACTGTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 824
Db 661 TGTGGAACACTGTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
QY 825 ACTGTGGGCTGTACAGCTTCTTCTGCACTTGTCTAGTTGGGCGGAGTTTCTGAACCA 884
Db 721 ACTGTGGGCTGTACAGCTTCTTCTGCACTTGTCTAGTTGGGCGGAGTTTCTGAACCA 780
QY 885 GCAAGGCTTACCTTGGGCGATGAGCTGAGCTGTGTGCGCTTCAAGTCTTCGAG 944
Db 781 GCAAGGCTTACCTTGGGCGATGAGCTGAGCTGTGTGCGCTTCAAGTCTTCGAG 840
QY 945 TTCTTTCTTATGTTGGCTGAGGCTGAGGCTTCTCAAGGCGCTGCTGGGCTGGAAG 1004
Db 841 TTCTTTCTTATGTTGGCTGAGGCTTCTCAAGGCGCTGCTGGGCTGGAAG 866
QY 1005 CATGGCCAGAGGGGTCAATGGCCAGACAGCTGTTGAGACGAGATGCAAGTGTACAGAAAG 1064
Db 867 ----- 866
QY 1065 AAGGCTTCAAGGCTGAGAAAGCAGCAGGCGTGTGGCGCAACCTGTAAATCCAGCTACT 1124
Db 867 ----- 866
```

```
QY 1125 CCGAGGCTGAGGAGAGAAATCGCTTGACCCCGGAGGCGAGAGTTGTGTGGCAGAGC 1184
Db 867 ----- 867
QY 1185 AGCTATGAACCCCTTGGAGAGATGATGATGATTTTGGAGCAACATGATTTGACA 1244
Db 878 AGCTATGAACCCCTTGGAGAGATGATGATGATTTTGGAGCAACATGATTTGACA 937
QY 1245 GGAATTTGACAGTGTCCCTGTGTGCTGTGATGATGATGATGATGATGATGATGATG 1304
Db 938 GGAATTTGACAGTGTCCCTGTGTGCTGTGATGATGATGATGATGATGATGATGATG 997
QY 1305 AGCCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1364
Db 998 AGCCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1057
QY 1365 AGTTCCGTCAGCTCTCTTATGAGGCTTCACTTCAACATGATGATGATGATGATGATG 1424
Db 1058 AGTTCCGTCAGCTCTCTTATGAGGCTTCACTTCAACATGATGATGATGATGATGATG 1117
QY 1425 TGAAGTTCCAGCCCAATCAGAGAGAGAGAGATGCTCAGCTGATCATTTGCGCT 1484
Db 1118 TGAAGTTCCAGCCCAATCAGAGAGAGAGAGATGCTCAGCTGATCATTTGCGCT 1177
QY 1485 TCTTAGGCTGCAAGTCCCATGATCACTATCTTCCAGGCGAACTCAAGGCAACATG 1544
Db 1178 TCTTAGGCTGCAAGTCCCATGATCACTATCTTCCAGGCGAACTCAAGGCAACATG 1237
QY 1545 TGTGGCCCAAGAGGATCCCTTCTCCAGAGGCGCTGCGCAAAACCAAGGCGCA 1604
Db 1238 TGTGGCCCAAGAGGATCCCTTCTCCAGAGGCGCTGCGCAAAACCAAGGCGCA 1297
QY 1605 AACAAGACCTTAAAGGCGCCAGAGAAACAACAAGGCTTGAAGCTTGAAGCGCT 1664
Db 1298 AACAAGACCTTAAAGGCGCCAGAGAAACAACAAGGCTTGAAGCTTGAAGCGCT 1357
QY 1665 TCAAGTCTGCGCCCACTGATCAGAGGCGAGCTACAGTCCAGGCGCCCAAGAGCGCC 1724
Db 1358 TCAAGTCTGCGCCCACTGATCAGAGGCGAGCTACAGTCCAGGCGCCCAAGAGCGCC 1417
QY 1725 GCGCCACTCCCATGTTCTTCCCTTGAAGCACTCAGCGCGTCAAGGCTTCAAGTGA 1784
Db 1418 GCGCCACTCCCATGTTCTTCCCTTGAAGCACTCAGCGCGTCAAGGCTTCAAGTGA 1477
QY 1785 CAGGCAATGACACCAAGCAAAAGCTTAAAGCTGATGATGATGATGATGATGATGATG 1844
Db 1478 CAGGCAATGACACCAAGCAAAAGCTTAAAGCTGATGATGATGATGATGATGATGATG 1537
QY 1845 TGAATTTGCTCTCAGAGAGCGATGAGGCGCTTGAAGGACCAAGGATATCTCAAGTGA 1904
Db 1538 TGAATTTGCTCTCAGAGAGCGATGAGGCGCTTGAAGGACCAAGGATATCTCAAGTGA 1597
QY 1905 GGAAGAAAACCTGTGAGATTTAACTGACGATATGCGAGATGCCGAAATATCACTCA 1964
Db 1598 GGAAGAAAACCTGTGAGATTTAACTGACGATATGCGAGATGCCGAAATATCACTCA 1657
QY 1965 AAGAACCTTTGGAGCAATCAACCAACATACATACATCACTCAAGGATGATGATGATG 2024
Db 1658 AAGAACCTTTGGAGCAATCAACCAACATACATACATCACTCAAGGATGATGATGATG 1717
QY 2025 CTTATTTGGGCTTGGAAAAAGGAGATGAGCAATTCCTAA 2065
Db 1718 CTTATTTGGGCTTGGAAAAAGGAGATGAGCAATTCCTAA 1758
```

RESULT 10
AX745964
LOCUS AX745964 1758 bp DNA linear PAT 14-MAY-2003
DEFINITION Sequence 1 from Patent WO03030922.
ACCESSION AX745964
VERSION AX745964.1 GI:30724619
KEYWORDS
SOURCE Homo sapiens (human)

ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.				
AUTHORS	1 Stuenkel, A., Bruenner, G., Fritsch, R., Eulenberger, K. and Ciosek, I.				
TITLE	Beatrophilin and beatrophilin homologous proteins involved in the regulation of energy homeostasis				
JOURNAL	Patent: WO 03030922-A 1 17-Apr-2003; Deutscher Patentamt; Entwicklungsbereich fuer Entwicklungsbiochemische Forschung (DE)				
FEATURES	Location/Qualifiers				
Source	1..1758				
	/organism="Homo sapiens"				
	/mol_type="Genomic DNA"				
	/db_xref="taxon:9606"				
ORIGIN					
Query Match	63.5%	Score 1541.8	DB 6	Length 1758	
Best Local Similarity	89.5%	Pred. No. 0			
Matches 1756	Conservative 0	Mismatches 2	Indels 203	Gaps 1	
Qy	105	ATGACCATCTTACACAAAGCAAGTGGCTATAGCCGCTTCCCTTCCCGCTG	164		
Db	1	ATGACCATCTTACACAAAGCAAGTGGCTATAGCCGCTTCCCTTCCCGCTG	60		
Qy	165	CTGCTGTCGTCGCGGCGAGCATCTCAAGCTGCTATATGGGAGTTCTTATCTTCTG	224		
Db	61	CTGCTGTCGTCGCGGCGAGCATCTCAAGCTGCTATATGGGAGTTCTTATCTTCTG	120		
Qy	225	CTGCTGTCATCATCATCCGCTTATATATAGGCTGCGCTTCCGGAAGAACAGCTG	284		
Db	121	CTGCTGTCATCATCATCCGCTTATATATAGGCTGCGCTTCCGGAAGAACAGCTG	180		
Qy	285	ATGTTTGAAGAACTGATCTGTATTTGGACAGCTACATCCAGCTATCCCATTTCTTC	344		
Db	181	ATGTTTGAAGAACTGATCTGTATTTGGACAGCTACATCCAGCTATCCCATTTCTTC	240		
Qy	345	GTGCTGGGCTTCTATCGTAGCGCTGTGTGACCCGCTGTGGAAACAGTACGAGAACCTG	404		
Db	241	GTGCTGGGCTTCTATCGTAGCGCTGTGTGACCCGCTGTGGAAACAGTACGAGAACCTG	300		
Qy	405	CCGTGGCCCGACCGCTCTCATAGCTGTGTGCGGCTTCTGTGAAGGCAAGGACGAGCA	464		
Db	301	CCGTGGCCCGACCGCTCTCATAGCTGTGTGCGGCTTCTGTGAAGGCAAGGACGAGCA	360		
Qy	465	GGCGGCTGTCGCGGCGAGCTCATCCGCTACGCTTCTGTGGGCAAGCTGCTATCTG	524		
Db	361	GGCGGCTGTCGCGGCGAGCTCATCCGCTACGCTTCTGTGGGCAAGCTGCTATCTG	420		
Qy	525	CGCAGCGTCAGACCGCAGTCTACAAAGGCTTCCCGAGCGCCAGCACTGTGTGAAGCA	584		
Db	421	CGCAGCGTCAGACCGCAGTCTACAAAGGCTTCCCGAGCGCCAGCACTGTGTGAAGCA	480		
Qy	585	GGCTTTATGACTCCGCGAGAACACAAAGAGTTGGAAATGAGCTTACCAACAATG	644		
Db	481	GGCTTTATGACTCCGCGAGAACACAAAGAGTTGGAAATGAGCTTACCAACAATG	540		
Qy	645	TTCTGGGTCCTCTGGGTCGTGTGCAACTGTCTATGAAGCGTGTCTTGAAGTGA	704		
Db	541	TTCTGGGTCCTCTGGGTCGTGTGCAACTGTCTATGAAGCGTGTCTTGAAGTGA	600		
Qy	705	ATCCGGGACCTTATCCGCTCAGAGCTGTGAACGAGATGAACAATTGCGATCTAG	764		
Db	601	ATCCGGGACCTTATCCGCTCAGAGCTGTGAACGAGATGAACAATTGCGATCTAG	660		
Qy	765	TGTGAGACCTGTATGCTTACGATGATTTAGTATCCCACTGTGTATACACAGTGTG	824		
Db	661	TGTGAGACCTGTATGCTTACGATGATTTAGTATCCCACTGTGTATACACAGTGTG	720		
Qy	825	ACTGTGGGCTGTACAGCTTCTTCTGACTTGTCTAGTTGGGCGGCAATTTCTGACCA	884		
Db	721	ACTGTGGGCTGTACAGCTTCTTCTGACTTGTCTAGTTGGGCGGCAATTTCTGACCA	780		

Qy	885	GCCAGGCTTACCTTGGCCATGAGCTGACCTGTTGGCCGCTTCTACAGTTCTTGGAG	944		
Db	781	GCCAGGCTTACCTTGGCCATGAGCTGACCTGTTGGCCGCTTCTACAGTTCTTGGAG	840		
Qy	945	TTCTTCTTCTATGTTGGCTGCTGAAGTGGGCTTCTCCAGGCGCTGCTGGGCTGAGG	1004		
Db	841	TTCTTCTTCTATGTTGGCTGCTGAGTGGGCTTCTCCAGGCGCTGCTGGGCTGAGG	866		
Qy	1005	CATGGCCAGAGGGGTGATGCGCAGAGCTGCTTGAAGAGAGATGACGTTCAGGAAGG	1064		
Db	867	-----	866		
Qy	1065	AAGGTCTACCGGATGAAAGAGCAGCGGTGTGGCGCAGACCTGTATCCAGCTACT	1124		
Db	867	-----	866		
Qy	1125	CGGAGGCTGAGGAGAGAAATCGCTTGAACCCCGGAGCGGAGGTTGTGTGCGAGG	1184		
Db	867	-----	866		
Qy	1185	AGCTCATCAACCCCTTGGAGAGATGATGATTTTGAACCACTGATTTGTCACA	1244		
Db	878	AGCTCATCAACCCCTTGGAGAGATGATGATTTTGAACCACTGATTTGTCACA	937		
Qy	1245	GGAATTTGCAAGTGTCCCTGTTGGCTGTGATGATGATGATGATGATGATGATGATG	1304		
Db	938	GGAATTTGCAAGTGTCCCTGTTGGCTGTGATGATGATGATGATGATGATGATGATG	997		
Qy	1305	AGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1364		
Db	998	AGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1057		
Qy	1365	AGTTCCGTCGAGCTTCTTATGAGCTTCACTTCAATATGAGCTTGAACCAAGAGGA	1424		
Db	1058	AGTTCCGTCGAGCTTCTTATGAGCTTCACTTCAATATGAGCTTGAACCAAGAGGA	1117		
Qy	1425	TGAGATTTCAGCCCAATCAGAGAGAGAGAGATGCTCAGCTGAGCATCTTGTGCGCT	1484		
Db	1118	TGAGATTTCAGCCCAATCAGAGAGAGAGAGATGCTCAGCTGAGCATCTTGTGCGCT	1177		
Qy	1485	TCCTAGGCTGCACTGCTTATGAGCTTCACTTCCAGGCGCAACTCAAGGACCAATAC	1544		
Db	1178	TCCTAGGCTGCACTGCTTATGAGCTTCACTTCCAGGCGCAACTCAAGGACCAATAC	1237		
Qy	1545	TGTGGCCCAAGAGGGAATCCCTTCTTCCAGAGGCTGTGCTTCCAAACCAAGGACCA	1604		
Db	1238	TGTGGCCCAAGAGGGAATCCCTTCTTCCAGAGGCTGTGCTTCCAAACCAAGGACCA	1297		
Qy	1605	AACAGAACTTAAAGGAGGCAAGAGAAACAAGAGGCTGGAAGCTTAAAGCTGTGAGCGCT	1664		
Db	1298	AACAGAACTTAAAGGAGGCAAGAGAAACAAGAGGCTGGAAGCTTAAAGCTGTGAGCGCT	1357		
Qy	1665	TCAAGTCTGGCCCACTGATCAGAGGCGAGCTTCACTAGTCCCAAGAGCGCCCTCA	1724		
Db	1358	TCAAGTCTGGCCCACTGATCAGAGGCGAGCTTCACTAGTCCCAAGAGCGCCCTCA	1417		
Qy	1725	GGCCCACTCCCATGTTCTTCCCTTGAACCATCAGAGGCGGTCAAGCTTCAAGTGTCA	1784		
Db	1418	GGCCCACTCCCATGTTCTTCCCTTGAACCATCAGAGGCGGTCAAGCTTCAAGTGTCA	1477		
Qy	1785	CAGGCAATGACACAAAGAGCAAAAGCTTAAAGCTGTGAGTTCTGGGGCCAAAGAAAT	1844		
Db	1478	CAGGCAATGACACAAAGAGCAAAAGCTTAAAGCTGTGAGTTCTGGGGCCAAAGAAAT	1537		
Qy	1845	TTGAATTGCTCTCAGAGAGGAGTGGGCTTGTATGAGACCAAGAGTATCTCAAGTGA	1904		
Db	1538	TTGAATTGCTCTCAGAGAGGAGTGGGCTTGTATGAGACCAAGAGTATCTCAAGTGA	1597		
Qy	1905	GGAGGAAACCTGTGAGTTTAACTGACGATATGCAAGAGATCCCGAAATCACTCA	1964		
Db	1598	GGAGGAAACCTGTGAGTTTAACTGACGATATGCAAGAGATCCCGAAATCACTCA	1657		

QY	1965	AGAACCTTTGGAAACATATCCCAACCAATATCACTACCTCAAGATCATGATGC	2024
Db	1658	AGAACTTTTGGAAACATATCCCAACCAATATCACTACCTCAAGATCATGATGC	1717
QY	2025	CTTATTGGGCTTTGGAAACAGGATGTAAGCATTCCTCA	2065
Db	1718	CTTATTGGGCTTTGGAAACAGGATGTAAGCATTCCTCA	1758
RESULT 11			
BC041664			
LOCUS			
DEFINITION	BC041664	2435 bp	mRNA
ACCESSION	BC041664.1	GI:27371319	
VERSION	MGCL		
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 2435)		
AUTHORS	Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klusner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Alechul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stopleton, M., Soares, M.B., Donald, M.F., Casavant, T.L., Scheet, T.E., Brownstein, M.J., Ueding, T.B., Tothlyuk, S., Mancini, P., Prange, C., Raha, S.S., Loggellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Boeck, S.A., McMan, P.U., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.W., Wolley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulys, S.W., Villalón, D.C., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kellerman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shewchenko, Y., Bonfield, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Gilmwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Small, D.E., Scherch, A., Schein, J.E., Jones, S.J., and Marra, M.A.		
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A.	99 (26)	16899-16903 (2002)
MEDLINE	22388257		
PUBMED	12477932		
REFERENCE	2 (bases 1 to 2435)		
AUTHORS	Strausberg, R.		
TITLE	Direct Submission		
JOURNAL	Submitted (20-DEC-2002)		
MEDLINE	Genetic Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
PUBMED			
REFERENCE	NIH-MGC Project URL: http://mgc.ncl.nih.gov		
AUTHORS	Contact: MGC help desk		
TITLE	Email: cgabs-remail.nih.gov		
JOURNAL	Tissue Procurement: Life Technologies, Inc.		
MEDLINE	cDNA Library Preparation: Life Technologies, Inc.		
PUBMED	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)		
REFERENCE	DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),		
AUTHORS	Gaithersburg, Maryland;		
TITLE	Web site: http://www.nisc.nih.gov/		
JOURNAL	Contact: nisc.mgc@nih.gov		
MEDLINE	Akheri, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaspy, J., Maduro, O.L., Masello, C., Maskeri, B., Mestrian, S.D., McCloskey, J.C., McDowell, C., Pearson, R., Stantip, S., Thomas, P.J., Touchman, J.W., Tseng, J.C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.		
COMMENT			

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLNC at: <http://image.llnl.gov>
 Series: IRAX plate: 82 Row: d Column: 4.
 Location/Qualifiers
 1..2435
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:47884 IMAGE:5194649"
 /tissue_type="Brain, adult, 6 pooled whole brains"
 /clone_lib="NH MGC_114"
 /lab_host="DH10B"
 /notes="vector: pCMV-SPORT6"
 1..2435
 /gene="VMD2"
 /notes="synonyms: BMD, BEST"
 /db_xref="locusid:7439"
 /db_xref="MIM:153700"
 87..1901
 /gene="VMD2"
 /codon_start=1
 /product="VMD2 protein"
 /protein_id="AA041664.1"
 /db_xref="GI:27371320"
 /db_xref="locusid:7439"
 /db_xref="MIM:153700"
 /translation="MEKRLTYLQDSYIOLIPISFVLGFVYVTLVTMRMNQVNP.BMPD
 RLMSLGFVFGKDEGRLLRLIRYANLGNVLLRSSTAYVRFPSAQNLYDAGF
 MTPLENHQLKLSLPHNMTFPPWVFWFANLSKMLDGLRLDPLTLOSLLNNNTLTQ
 CGHLYADWLSIPLVYQVTVAVSFPLTGVGRQLNPAPAYRGHELDLVVPVTF
 LOFFVYGVMLKVAALQLINPEDDDDFETNMWIDNRLQVSLAVDEMQLDPRMPDM
 LYNNPEQPPYTAASQFRFVSFMGSLFNISLNKEEMEPQNOBEBDAHGIIGRFL
 GLQSHDHPPRANSRPTLMPKRESLHSGELPKNHKAQKNVRGQDNKMAKLVADA
 FKSPALYORPGYSAAPRTLPSPMPFPLEPSAPSKLSVTVGIDITPKDSIKTVSSGAK
 KSEPLISESGALMHPVSVQRKTYERNUTDMEPIPNHLKBEPLDSPTNITHTLK
 DHNDPYALENRSVLHNDQHCIALCPTASIALSLSPFLHNLGFRHQCSTLDLRPAL
 AMGIYLVTFGLIGKCSGPLTSPMYHPDFLQPSGR"
 90..1022
 /gene="VMD2"
 /notes="dup289; Region: Putative membrane protein. This
 family called family 0 in, may be a transmembrane protein
 The specific function of this protein is unknown"
 /db_xref="CDD:pfam01062"

```
Qy 617 GAGAACTGAGCCTACACAA CATGTTCTGGGTGCTTGGTGTGTTGCCAACT 676
Db 419 GAGAACTGAGCCTACACAA CATGTTCTGGGTGCTTGGTGTGTTGCCAACT 478
Qy 677 GTCAATGAGCGGTGGTGGAGGTCCGATCCGGGACCTTATCTGCTCCAGAGCTGCT 736
Db 479 GTCAATGAGCGGTGGTGGAGGTCCGATCCGGGACCTTATCTGCTCCAGAGCTGCT 538
Qy 737 GAAAGATGAAACA CTTGGCTACTCACTGATGGA CACTGATGCTACGATGATG 796
Db 539 GAAAGATGAAACA CTTGGCTACTCACTGATGGA CACTGATGCTACGATGATG 598
Qy 797 TATCCACTGTGTATACACAGTGTGACTGTGGCGTGTATACAGCTTCTCTGACTTG 856
Db 599 TATCCACTGTGTATACACAGTGTGACTGTGGCGTGTATACAGCTTCTCTGACTTG 658
Qy 857 TCTAGTTGGGCGGAGTTTCTGA ACCGACCAAGGCTTACCTTGCCATAGCTGAGCT 916
Db 659 TCTAGTTGGGCGGAGTTTCTGA ACCGACCAAGGCTTACCTTGCCATAGCTGAGCT 718
Qy 917 CGTTGTGCGGCTCTTCA GGTCTCTGAGTCTTCTATGTTGGCTGGAGAGTGG 976
Db 719 CGTTGTGCGGCTCTTCA GGTCTCTGAGTCTTCTATGTTGGCTGGAGAGTGG 772
Qy 977 CCTCTCAGGCGCTGTGCTGGGCTGGAAGCATGACCAGAGGGGTATGGCCAGACTGCT 1036
Db 773 ----- 772
Qy 1037 TGAGACGAGATGATGATGTCAGAAAGAGTCTCACGGGTAGAAAGACGACGCGCTG 1096
Db 773 ----- 772
Qy 1097 GTGGCGCAACCTGTATCCAGCTACTCGGGAGGCTGAGCAGAGAAATGCTTGAACC 1156
Db 773 ----- 772
Qy 1157 CGGAGCGGAGGTTGTGTGGCAGAGAGCTCATCAACCCCTTGGAGAGATGATGAT 1216
Db 773 ----- 815
Qy 1217 GATTTGAGACCACTGATTTGTGACAGAAATTTGACAGGTCTCCCTGTTGGCTGTGAT 1276
Db 816 GATTTGAGACCACTGATTTGTGACAGAAATTTGACAGGTCTCCCTGTTGGCTGTGAT 875
Qy 1277 GAGATGACCAAGGACCTGCTCGGATGGAAGCCGACATGTACTGGAATTAACCCGAGCA 1336
Db 876 GAGATGACCAAGGACCTGCTCGGATGGAAGCCGACATGTACTGGAATTAACCCGAGCA 935
Qy 1337 CAGCCCCCTTACAGAGCTGCTTCGCGCAGTTCGCTGAGGCTCTTATATGGGCTCCACC 1396
Db 936 CAGCCCCCTTACAGAGCTGCTTCGCGCAGTTCGCTGAGGCTCTTATATGGGCTCCACC 995
Qy 1397 TTCAACATCAGCCTGAAACAAAGAGATGAGATTCCAGCCCAATCAGAGGAGCAGAGAG 1456
Db 996 TTCAACATCAGCCTGAAACAAAGAGATGAGATTCCAGCCCAATCAGAGGAGCAGAGAG 1055
Qy 1457 GATCTCAAGCTGGCATATTTGGCCGCTTCTTAGGCTGCGAGTCCCATATCACCATCTCT 1516
Db 1056 GATCTCAAGCTGGCATATTTGGCCGCTTCTTAGGCTGCGAGTCCCATATCACCATCTCT 1115
Qy 1517 CCCAGGCAAACTAAGAGCAAACTACTGTGGCCCAAGAGGAAATCCCTTCTCAGAG 1576
Db 1116 CCCAGGCAAACTAAGAGCAAACTACTGTGGCCCAAGAGGAAATCCCTTCTCAGAG 1175
Qy 1577 GAGCTGCCCCAAAACCAACAGGACGCAACAGACGTTAGAGGCGCAGACCAAG 1636
Db 1176 GAGCTGCCCCAAAACCAACAGGACGCAACAGACGTTAGAGGCGCAGACCAAG 1235
Qy 1637 GCGCTGAGAGCTTAAAGCTGTGAGAGCTTCAAGTGTGCGCCATCTGTATCAGAGCCAGGC 1696
Db 1236 GCGCTGAGAGCTTAAAGCTGTGAGAGCTTCAAGTGTGCGCCATCTGTATCAGAGCCAGGC 1295
Qy 1697 TACTACAGTGGCCACAGAGCGCCCTCAGCCCACTCCATGTTCTTCCCTTGAACA 1756
```

```
Db 1296 TACTACAGTGGCCACAGAGCGCCCTCAGGCCCACTCCCATGTTCTTCCCTTGAACA 1355
Qy 1357 TCAGCGCCCTTCAAGCTTCAACAGTGTCAACAGGATTAAGACCAAGAAAGCTTAAAG 1816
Db 1356 TCAGCGCCCTTCAAGCTTCAACAGTGTCAACAGGATTAAGACCAAGAAAGCTTAAAG 1415
Qy 1817 ACTGTAGTTCTGTGGGCGCAAGAAAGTTTGAATTGCTCTCAGAGAGGATGGGCGCTTG 1876
Db 1416 ACTGTAGTTCTGTGGGCGCAAGAAAGTTTGAATTGCTCTCAGAGAGGATGGGCGCTTG 1475
Qy 1877 ATGAGACACCCAGAAAGTATCTCAAGTGAAGAGAAACCTGTGAGTTTAACCTGAGAT 1936
Db 1476 ATGAGACACCCAGAAAGTATCTCAAGTGAAGAGAAACCTGTGAGTTTAACCTGAGAT 1535
Qy 1937 ATGCCAAGATCCCGGAAATACCTCAAGAAACCTTTGGAAACATACCAACCAACTA 1996
Db 1536 ATGCCAAGATCCCGGAAATACCTCAAGAAACCTTTGGAAACATACCAACCAACTA 1595
Qy 1997 CACACTACACTCAAAAGATCATGATGATCCTTATTGGGCGCTTGAAGAAACAGGATG 2051
Db 1596 CACACTACACTCAAAAGATCATGATGATCCTTATTGGGCGCTTGAAGAAACAGGATG 1650

RESULT 12
BD132659 1263 bp DNA Linear PAT 18-SEP-2002
LOCUS
DEFINITION Secreted proteins and polynucleotides encoding them.
ACCESSION BD132659
VERSION BD132659.1 GI:23227604
KEYWORDS JP 2002504822-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Jacobs,K., McCoy,J.M., Lavallie,B.R., Racie,L.A., Treac,M.,
Spaulding,V., Agostino,M.J., Howes,S.H. and Reachel,K.
TITLE Secreted proteins and polynucleotides encoding them
JOURNAL Patient: JP 2002504822-A 2 12-FEB-2002;
GENETICS INSTITUTE INC
PN JP 2002504822-A/2
PD 12-FEB-2002
PF 08-JUN-1998 JP 1999503038
PR 11-JUN-1997 US 08/873218, 05-JUN-1998 US 09/092722 PI
KENNETH JACOBS,JOHN M MCCOY,EDWARD R LAVALLIE,LISA A RACIE,PI
MURICE TREACY,
PI VIKKI SPAULDING,MICHAEL J AGOSTINO,STEVEN H HOWES,KIM PECHTEL
PC C12N15/12,C07K14/47,A61K38/17
CC Strandedness: Double;
CC Topology: Linear;
FH Key Location/Qualifiers.
FEATURES
source location/Qualifiers
1..1263
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 51.2%; Score 1244.2; DB 6; Length 1263;
Best Local Similarity 99.8%; Pred. No. 2.2e-284;
Matches 1246; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1181 GAGCAGCTCATCAACCCCTTTGAGAGAGATGATATTTTGGACCAACTGATTTGTC 1240
Db 1 GAGCAGCTCATCAACCCCTTTGAGAGAGATGATATTTTGGACCAACTGATTTGTC 60
Qy 1241 GACGAGATTTCAGGTGCTCCCTTGGCTGTGATGAGATGACACGAGACCTGCTCGG 1300
Db 61 GACGAGATTTCAGGTGCTCCCTTGGCTGTGATGAGATGACACGAGACCTGCTCGG 120
Qy 1301 ATGAGCGGAGCATGTACTGGAATTAAGCCGAGCCAGAGCCCTTACACAGCTGCTTCC 1360
```

```

Db      121 ATGAGCGCGA CATGTA CTGTAATG AAGCCG GAGCCAG ACCCCCTT AGACA GCTGCTCC 180
Oy      1361 GCCAGTTC CGTCGAG GCGCTC CTCTTA TGGGCTC CCACTT CAACAT AGCTGA ACAAAG 1420
Db      181 GCCCGATTC CGTCGAG GCGCTC CTCTTA TGGGCTC CCACTT CAACAT AGCTGA ACAAAG 240
Oy      1421 GAGATGAG GTTCCAG CCGCAAT CAGAGAG ACGAGAG ATGCTC CACGCT GGCATC ATTGGC 1480
Db      241 GAGATGAG GTTCCAG CCGCAAT CAGAGAG ACGAGAG ATGCTC CACGCT GGCATC ATTGGC 300
Oy      1481 CGCTTCTA AGCTCCTG CACTCC CATATAT CCACTT CCGAGG GCAAACT CAAAGACCAA 1540
Db      301 CGCTTCTA AGCTCCTG CACTCC CATATAT CCACTT CCGAGG GCAAACT CAAAGACCAA 360
Oy      1541 CTACTGTG GCGCCAA GAGGAAAT CCCTT CCAACG AGGCGCTG CCGCAAAA CCAAGGCA 1600
Db      361 CTACTGTG GCGCCAA GAGGAAAT CCCTT CCAACG AGGCGCTG CCGCAAAA CCAAGGCA 420
Oy      1601 GCCAAACG AAGCTTA GGGGCG CAGGAG ACAAACA AAGCGCTG GAAAGCTT AAGGCTGTGAC 1660
Db      421 GCCAAACG AAGCTTA GGGGCG CAGGAG ACAAACA AAGCGCTG GAAAGCTT AAGGCTGTGAC 480
Oy      1661 GCGTTCAAG TCTGGCC CACTGT ATATAG AGGCGCTC TCAAGT GCGCCCA CAGACGCCC 1720
Db      481 GCGTTCAAG TCTGGCC CACTGT ATATAG AGGCGCTC TCAAGT GCGCCCA CAGACGCCC 540
Oy      1721 CTGAGCCCA CTCCCAT GTTCTT CCGCCCTA GAAACCAT CAGCGCGT CAAAGCTT CACAGT 1780
Db      541 CTGAGCCCA CTCCCAT GTTCTT CCGCCCTA GAAACCAT CAGCGCGT CAAAGCTT CACAGT 600
Oy      1781 GTCAAGGCA TATGAC CCAAAAGA CAAAGCTT AAAGA CTGAGT TCGGGGCCA AGAAA 1840
Db      601 GTCAAGGCA TATGAC CCAAAAGA CAAAGCTT AAAGA CTGAGT TCGGGGCCA AGAAA 660
Oy      1841 AGTTTGAAT GTGCTC TGAAGA GAGGAT GGGGCTT GATGAG GCAACCCG AAGTATCT CAA 1900
Db      661 AGTTTGAAT GTGCTC TGAAGA GAGGAT GGGGCTT GATGAG GCAACCCG AAGTATCT CAA 720
Oy      1901 GTGAGAGG AAGAACTG TGAATTA CTTGAC GGAATATG CAGAGAT CCGCGA AATCAC 1960
Db      721 GTGAGAGG AAGAACTG TGAATTA CTTGAC GGAATATG CAGAGAT CCGCGA AATCAC 780
Oy      1961 CTCAAGAA CCTTTGAA ACAAATC CAAACA CAACTA CACTAC AACTCA AATCAGATG 2020
Db      781 CTCAAGAA CCTTTGAA ACAAATC CAAACA CAACTA CACTAC AACTCA AATCAGATG 840
Oy      2021 GATCCTTAT TGGGCTT GGAAGAA CAGGAT AATGACA CACTTCTT AACTGCTT CTAATGG 2080
Db      841 GATCCTTAT TGGGCTT GGAAGAA CAGGAT AATGACA CACTTCTT AACTGCTT CTAATGG 900
Oy      2081 GGATGCTT GCGCAG CAGGCT CCACTG TGTGTA CCAAGAG ACACTGAT CCACTCA 2140
Db      901 GGATGCTT GCGCAG CAGGCT CCACTG TGTGTA CCAAGAG ACACTGAT CCACTCA 960
Oy      2141 CAGGCATTA CAGCTGT CCAACTG AAGAA CAGTGTCTT ACAAAG CCGTGAATCAAATGCTTA 2200
Db      961 CAGGCATTA CAGCTGT CCAACTG AAGAA CAGTGTCTT ACAAAG CCGTGAATCAAATGCTTA 1020
Oy      2201 GCTTAATGA TAAATAAAT CCGAGACTT AACTTCA GCGCTTAA TGCCCTTATTA TCAATAAACT 2260
Db      1021 GCTTAATGA TAAATAAAT CCGAGACTT AACTTCA GCGCTTAA TGCCCTTATTA TCAATAAACT 1080
Oy      2261 GTGAAAGCTA GACTGA ACAAATG GAAACATTTA ACTGAGA CTGGAAT CAGAGT CCGGGA 2320
Db      1081 GTGAAAGCTA GACTGA ACAAATG GAAACATTTA ACTGAGA CTGGAAT CAGAGT CCGGGA 1140
Oy      2321 ACCCTTAT GTTCTAT CTGAATC CAAAGACCA CACTTA GATAT CTGCCCAA ACTAATGA 2380
Db      1141 ACCCTTAT GTTCTAT CTGAATC CAAAGACCA CACTTA GATAT CTGCCCAA ACTAATGA 1200
Oy      2381 GTTTAATAA TPAACAAT TCTCGTTA AAAAAAAAAA AAAAAAAAAA 2429
Db      1201 GTTTAATAA TPAACAAT TCTCGTTA AAAAAAAAAA AAAAAAAAAA 1249

```

```

RESULT 13
BC079048
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 2035)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shennan,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Datchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Adamsom,R.D., Muliahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Morley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,A.C., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalske,U., Smallus,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 2035)
Director MGC Project.
Direct Submission
Submitted (02-AUG-2004) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.ncl.nih.gov
Contact: MGC help desk
Email: cgadps@email.nih.gov
Tissue Procurement: Howard Jacobs
CDNA Library Preparation: Express Genomics
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/BLN ac: http://image.llnl.gov
Series: IRK Plate: 183 Row: a Column: 22
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, similarity but not identity to protein.
Location/Qualifiers
1..2035
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="MGC:93985 IMAGE:7115295"
/rissue_type="Testis, rat (Brown Norway)"
/clone_id="NIH_MGC_237"
/lab_host="DH10B"
/notes="Vector: pExpress1"

```

```

FEATURES
source

```


Db 1118 GCTCCACCTTCACATCAGCCTTAAGAAAGAAAGACTTGAAGCTTTGTGTCAAAAGAGAGG 1177
Qy 1449 ACAGAGAGATGCTCAGCCTG-----GCATCATTTGGCGGCTTCTAGGCGTCGC 1496
Db 1178 CTGACACGAGATTAAGAAAGAGAGTGGCTATAGCAGACCATAGAGCTGCTTTTAGGACGTGC 1237
Qy 1497 AGTCCCATGATCACCATCTCCAGGCGAAACTCAGAGACCAAACTACTGTGGCCCAAGA 1556
Db 1238 AACCCAAAACATACATCTTCCCTGAAAGACTTAAAGACCAAACTATTGTGTTCTAAGA 1297
Qy 1557 GGGAAATCCCTTCTCAGCAGGCGCTGCCCAAAAACCAAGGACGCAACAGAAAGCTTA 1616
Db 1298 -----ACCCCTCTCCAGAGCCAGTGTAAAGATGCCAACAGAAAACAGAAAG--- 1348
Qy 1617 GGGGACAGAGAGACAAAGGCGTGGAGCTTAAAGGCTGTGAGCGCTTCAAGTCTGGGC 1676
Db 1349 -----ATGCTGGAAATTTAAAGGCTGTGAGCTTCTTGAATGTGTTC 1390
Qy 1677 CACTGTATCAGAGGCGCAGCTACTACAGTGCCTCCACAGACGCGCCCTCAGCCCACTCC 1734
Db 1391 CAAGTTTAAAGAGAGAGGCTCCCATTTGTGGCCCAAGAGCAAGCCAGACCACTTAC 1448

RESULT 15
BD136744 1916 bp DNA linear PAT 18-SEP-2002
LOCUS Beethoven's macular dystrophy gene.
ACCESSION BD136744
VERSION BD136744.1 GI:23231689
KEYWORDS JP 2002504559-A/26.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 1916)
AUTHORS Petrunkhin, K., Caskey, T. C., Metzker, M. and Wadelius, C.
TITLE Beethoven's macular dystrophy gene
JOURNAL Patent: JP 2002504559-A 26 12-FEB-2002;
MERCK & CO INC, CLABS WADELIVUS
COMMENT OS Mus musculus (mouse)
PN JP 2002504559-A/26
PD 12-FEB-2002
PR 22-FEB-1999 JP 2000533447
PR 25-FEB-1998 US 60/075941, 18-DEC-1998 US 60/112926 PI
KONSTANTIN PETRUNKHIN, THOMAS C CASKEY, MICHAEL METZKER, CLABS PI
WADELIVUS
PC C07K16/18, C07K14/47, C12N5/10, C12N5/09, C12P19/34, C12Q1/68// PC
C12P21/08,
CC C12N5/00, C12N5/00
CC Beethoven's macular dystrophy gene
FH key Location/Qualifiers
FT source 1.1916 Location/Qualifiers
FT Location/Qualifiers
1.1916
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"

ORIGIN
Query Match 27.1%; Score 658.2; DB 6; Length 1916;
Best Local Similarity 67.6%; Pred. No. 3.5e-145; Indels 242; Gaps 4;
Matches 1116; Conservative 0; Mismatches 293;

Qy 96 TGCCGTGCGATGACATCACTTACACAGCCAGTGTGCTTAATGCCGCTTAGGCTCTTC 155
Db 2 TGCCAGAGCATGATCACTACCAACAAAGTAGCCATGCCGCTCGTTTCGTTTC 61
Qy 156 TCCGCGCTGCTGCTGTGGCGGGGAGCATCTAACAGCTGTATATG3GAGTTCTTA 215
Db 62 TCGTCCCTCTCTGCTGTGGCGAGGAGCATCTAACAGCTGTATATG3GAGAAATTCCTT 121
Qy 216 ATCTTCCTGCTCTGCTATACATATCGGCTTTATTTATAGGCTGGCCCTCAGGAGAA 275

Db 1122 GTCTTCATATTCCTCTCTATATTCATCCGTGACTCTACAAATGTTCTCTCAGTAT 181
Qy 276 CAACAGCATGATTTGAAACATGACTCTGATTTGGCAGACTACATCAGTATCC 335
Db 182 CAGAGCTGTGTGTGAAGAGCTGCTGTACTGCGACACTACATTCAGTATCTCC 241
Qy 336 ATTTCTTCTGCTGTGGCTTCTACGTGACGCTGTGTGACCCCTGTGTGAAACGATAC 395
Db 242 ATATCTTCTGTTCTGGGTTCTATGTATCATTTGATGTGAGACCCCTGTGTGAGCAGTAC 301
Qy 396 GAGAACTGCGGCTGCGCCGACCGCTTATGAGCTGTGTGCGCTTGTGAAAGCAG 455
Db 302 GAGAACTTGGCGGCGCCGACCGCTTATGAGCTGTGTGAGCTTGTGTGAGGCGCAG 361
Qy 456 GAGAGCAAGGCGGCTGCGGCGGACCGCTTATCCGCTACGCGCAACCTGCGGACGTTG 515
Db 362 GATGAGAAAGCCCTTTGCTGCGGCGACCGCTTATCCGCTACGCGCTGCGGCGCAAGT 421
Qy 516 CTCATCTGCGCAGCTCAGACCGCAGTCTACAGAGGCTTCCCGAGCGCCAGCAGCTG 575
Db 422 CTCATCTGCGCAGCTCAGACCGCAGTCTCAGAGCGCTTCCCATCTTCAACACCTG 481
Qy 576 GTGCAAGAGGCTTTATGACTCCGCGAGAACACAGAGCTTGAGAACTGAGCTTACCA 635
Db 482 GTGCTAGAGAGTTTATGACCCTAGGAGAACATAGCAGTTGACAGAGTTGGGCTTACA 541
Qy 636 CACAACTGTTCTGCGGCTGCGGCTGTGTGTGCTTCCAACTGTATGAAAGCGTGT 695
Db 542 CACAACTGTTCTGCGGCTGCGGCTGTGTGTGCTTCCAACTGTATGAAAGCGTGT 601
Qy 696 GAGAGTGAATCCGAGGACCCCTATCTGCTCCAGAGCGCTGTGAAGAGATGAACCTTG 755
Db 602 GAGAGTGAATCCGAGGACCCCTATCTGCTCCAGAGCGCTGTGAATGAGGTGTGACTTG 661
Qy 756 GGTACTGAGTGTGACACCTGTATACCTGACCTGATGATTAATACCACTGTGTATACA 815
Db 662 GGTACTGAGTGTGACACCTGTATACCTGACCTGATGATTAATACCACTGTGTATACA 721
Qy 816 CAGGTGTGACTGTGGCGGTGTACAGCTTCTCTGACTGTCTGATGTTGGCGGAGTTT 875
Db 722 CAGGTGTGACTGTGGCGGTGTACAGCTTCTCTGACTGTCTGATGTTGGCGGAGTTT 781
Qy 876 CTGAACCAAGCAGGCGCTACCTGCGCATAGCTGAGACCTGTTGGCGGCTTACAG 935
Db 782 CTGAACCAAGCAGGCGCTACCTGCGCATAGCTGAGACCTGTTGGCGGCTTACAG 841
Qy 936 TTCTGAGTCTTCTTCTATATGTTGCTGCTGAGAGGTGGGCTCTCCAGGCGGCTGCTG 995
Db 842 ATCTGCAATTTCTTATTTCTATATGGGCTGCTGAA----- 876
Qy 996 GGTGTGAGGCTGGCCAGAGGGGTATGCGCAGAGCTGCTTGAAGCAGGATGCACTGT 1055
Db 877 ----- 876
Qy 1056 CAGGAAAGAAAGTCTACAGGCTGAGAAAGCAGCCAGCGGTGTGGCCGACCTGTATC 1115
Db 877 ----- 876
Qy 1116 CCAGCTACTGGAGAGGCTGAGGAGAGAAATCGCTTGAACCGGAGGCGAGGTTGTGG 1175
Db 877 -----GG 878
Qy 1176 TGCGAGAGCAGCTGATCAACCCCTTTGAGAGAGATGATGATTTTGAACCACTGGA 1235
Db 879 TGCGAGAGCAGCTGATCAACCCCTTTGAGAGAGATGATGATTTTGAAGACTTAATCTGGA 938
Qy 1236 TTGTGACAGGAATTTGAGAGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1295
Db 939 TCATTTGACAGAAACCTGAGGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 998
Qy 1296 CTGAGATGAGGCGGAGCATGTATGAGATAAGCCGAGGCCAGCCGCCCTTACAGAGCTG 1355


```

Db      999  CTCCCATGGAACGTGACATGTACTGGAACGAGGCAAGCGCTCAGCCGCCCTACACAGCTG 1058
QY      1356  CTTCCGCGCCAGTTCGCTGCGAGCCCTCTTATGGGCTCCACCTTCAACATGAGCCTGAACA 1415
Db      1059  CTTCTGCGAGGCTCTGCGCGGATTCCTTCATGAGGCTCCACCTTCACATCAGCCTAAGA 1118
QY      1416  AAGAGAGATGAGGTTCAGGCCCAATCAGAGAGACGAGAGATGCTCACGCTG----- 1469
Db      1119  AAGAGAGCTTAGACCTTGTGCATAAAGAGAGGCTGACACGATTAAGAAAGAGGTGGCT 1178
QY      1470  -----GCATCATTTGGCCGCTTCTCTAGGCTTGCAATCCATGATCACCATCTCCAGGG 1523
Db      1179  ATAGCAGACCATAGGCTGCTTCTTAGGACTGCAACCCAAAACCTACCATCTTCCCTTGA 1238
QY      1524  CAAACTGAAGGACCAACTACTGTGGGCCCAAGAGGAAATCCCTTCTCCAGAGGGGCTGC 1583
Db      1239  AAGACTTAAAGACCAAACTATTGTGTTTAAAG-----ATCCCTCTCTCGAAGGCCAGT 1292
QY      1584  CCAAAAAACACAGGACAGCCAAACAGAACGTTAGGGGCCAGAAAGACAAACAGGCTTGA 1643
Db      1293  GTAAGATGCCAACCAAGAAAACAGAAAG-----ATGCTTGA 1331
QY      1644  AGCTTAAAGCTGTGAGCGCTTCAAGTCTGGCCCATGTATCAGAGGCCAGGCTACTACA 1703
Db      1332  AATTAAAGGCTGTGACTTCTTGAATGTGTCCAAAGGTTTAAAGAGAGAGGCTCCATT 1391
QY      1704  GTGCCCCACAGAGCGCCCTCAGCCCACTCC 1734
Db      1392  GTGGCCCAAGGCAACCCAGAGCCACCTTAC 1422

```

Search completed: March 26, 2005, 07:38:57
 Job time : 10382.6 secs

1	661.8	28.1	711.5	EX095540
2	657	27.0	676.5	BU741926
3	651.6	26.8	659.5	BU730894
4	640.8	26.7	674.5	BU731809
5	640.4	26.4	666.4	BI480798
6	638	26.3	639.6	CA389968
7	631.2	26.0	636.6	CA397981
8	626.6	25.8	651.5	BU731149
9	582.4	24.0	592.4	BM663028
10	577	23.8	573.2	BE410951
11	575	23.7	585.4	BM707948
12	565.2	23.3	593.2	BE385592
13	564.6	23.2	583.5	BO879680
14	553.2	22.8	561.6	CDS19875
15	551.6	22.7	756.7	CO396179
16	541.2	22.3	1659.3	AK006549
17	538.8	21.8	735.4	BI756248
18	525.8	21.6	537.4	BM691446
19	522	21.5	522.4	BM685396
20	507.4	20.9	531.4	BM685396
21	485.4	20.4	497.6	CA391418
22	482.4	20.3	930.5	BO438624
23	492.2	20.2	545.6	CA395098
24	491	20.2	491.7	CV024947
				CV024947_2492_Full

C	25	485.2	20.0	566	4	BM718338	UI-E-EOI-
C	25	476.4	19.6	582	5	BP379588	BP379588
C	27	475.4	19.6	503	4	BM685122	UI-E-EJ1-
C	28	475.4	19.6	508	4	BM932117	UI-E-EJ1-
C	29	470.8	19.6	474	1	AI190190	UI-E-EJ1-
C	30	469.6	19.3	484	5	BU726009	UI-E-C10-
C	31	467	19.2	492	5	BU741783	UI-E-EJ0-
C	32	464.8	19.1	469	1	AA307119	ES1718031
C	33	444.6	18.3	601	4	BG951790	MR1-CT073
C	34	443	18.2	481	4	BM625514	UI-E-C11-
C	35	427.4	17.6	484	7	N31453	YX55008_r1
C	36	422	17.4	448	5	BU733195	UI-E-C11-
C	37	419	17.2	525	5	BO345562	PM3-NT031
C	38	405.4	16.7	419	4	BM718146	UI-E-EJ0-
C	39	402.6	16.6	792	7	CK470840	AGENCYCOURT
C	40	391	16.1	766	2	BE275846	GO1121656
C	41	388	16.0	713	6	BY714331	BY714331
C	42	385.2	15.9	862	7	CK469826	AGENCYCOURT
C	43	379.8	15.6	1067	4	BM652042	BM652042
C	44	378.4	15.6	381	1	AI208843	GG38912_x
C	45	375.8	15.5	688	6	CA441116	UI-H-DE0-

ALIGNMENTS

RESULT_1	
LOCUS	BX095540
DEFINITION	711 bp mRNA linear EST 04-FEB-2003
ACCESSION	BX095540 Soares melanocyte 2NBHM Homo sapiens cDNA clone
VERSION	IMAGE998B16582 ; IMAGE:265647, mRNA sequence.
KEYWORDS	BX095540
SOURCE	BX095540.1 GI:27827758
ORGANISM	EST. Homo sapiens (human)
REFERENCE	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 711) Ebert,L., Hettl,O., Hennig,S., Neubert,P., Patsch,E., Peters,M., Radeloff,U., Schneider,D. and Korn,B. Human Unigeneset - RZPD3 Unpublished (2003)
TITLE	Contact: Ina Rolfs
JOURNAL	RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
COMMENT	RZPD; IMAGE998B16582. RZPDLIB; I.M.A.G.E. CDNA Clone Collection; Human Unigeneset - RZPD3 (RZPDLIB No.972) http://www.rzpd.de/CloneCard/cgi- bin/showlib.pl.cgi/response?libNo=972 Contact: Ina Rolfs RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH Heubnerweg 6, D-14059 Berlin, Germany Tel: +49 30 32639 101 Fax: +49 30 32639 111 www.rzpd.de This clone is available royalty-free from RZPD, contact RZPD (clone@rzpd.de) for information. Seq primer: MJ3r, Primer sequence: TTTCACACAGGAACGCTATGAC. Location/Qualifiers 1..711 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE998B16582 ; IMAGE:265647" /sex="Male" /rsluse_type="melanocyte" /lab_host="DH10B (ampicillin resistant)" /clone_lib="Soares melanocyte 2NBHM" /note="Vector: pT73D (Pharmacia) with a modified polylinker. Site_1: Not I, Site_2: Eco RI, 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCACACTGAACTGGAGCGCCGCAgTTTTTTTTTTTTTTTTTTT 3'] ,
FEATURES	
source	

double-stranded cDNA was size selected, ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library constructed by Bento Soares and M. Patricia Bonaldi. RNA from normal foreskin melanocytes (FS374) was kindly provided by Dr. Anthony P. Albino."

ORIGIN

Query Match 28.1%; Score 681.8; DB 5; Length 711;
Best Local Similarity 99.6%; Pred. No. 1.9e-137;
Matches 694; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

```

QY 1536 CCAACTACTGTGTGGCCCAAG-AGGGAATCCCTTCTCCAGAGGGCCCTGCCAAAACAC 1594
DB 1 CCAACTACTGTGTGGCCCAAGAGGAATCCCTTCTCCAGAGGGCCCTGCCAAAACAC 60
QY 1595 AAGGACGCCAAAGAAAGCTTAAAGGGCCGAGAAACAAGAGGCTTGAAGCTTAAAGCT 1654
DB 61 AAGGACGCCAAAGAAAGCTTAAAGGGCCGAGAAACAAGAGGCTTGAAGCTTAAAGCT 120
QY 1655 GTGGAAGCCTTCAAGTCTGGCCCACTGTATCAGAGGCCAGGCTACTAGAGTCCACAG 1714
DB 121 GTGGAAGCCTTCAAGTCTGGCCCACTGTATCAGAGGCCAGGCTACTAGAGTCCACAG 180
QY 1715 AGCCCTCTCAGCCCTCCATCTCTTCTCCCTAGAAACCATCAGCCGCTCAAAAGCTT 1774
DB 181 AGCCCTCTCAGCCCTCCATCTCTTCTCCCTAGAAACCATCAGCCGCTCAAAAGCTT 240
QY 1775 CACAGTGTACAGAGCATACACCAAGAAAGCTTAAAGCTGATGATTTCTGGGGCC 1834
DB 241 CACAGTGTACAGAGCATACACCAAGAAAGCTTAAAGCTGATGATTTCTGGGGCC 300
QY 1835 AAGAAAGTTTGAATTTGCTTCAAGAGCGATGGGGCCTTGATGAGACCCAGAAAGTA 1894
DB 301 AAGAAAGTTTGAATTTGCTTCAAGAGCGATGGGGCCTTGATGAGACCCAGAAAGTA 360
QY 1895 TCTCAAGTGAAGAGAAACCTGTGAGATTAACCTGACGAGATATCCAGAGATCCCGCA 1954
DB 361 TCTCAAGTGAAGAGAAACCTGTGAGATTAACCTGACGAGATATCCAGAGATCCCGCA 420
QY 1955 AATCACTTCAAGAACTTTGGACATACCAACCAATACACATACACATCACTAAGAT 2014
DB 421 AATCACTTCAAGAACTTTGGACATACCAACCAATACACATACACATCACTAAGAT 480
QY 2015 CACATGATCTTATTTGGGCTTGGAAAAACAGGATGAAGACATCTTCACTGCTTCC 2074
DB 481 CACATGATCTTATTTGGGCTTGGAAAAACAGGATGAAGACATCTTCACTGCTTCC 540
QY 2075 TAATGGGATGCTTGGCCAGCAGGTCTCTCACTGTGTGTACACAGCAGAGACTGATC 2134
DB 541 TAATGGGATGCTTGGCCAGCAGGTCTCTCACTGTGTGTACACAGCAGAGACTGATC 600
QY 2135 CAGTCACAGCATAAGCTGTCACTCACTGAAGAAAGTGTCTTCAAGAGCTGAATCAA 2194
DB 601 CAGTCACAGCATAAGCTGTCTCACTGAAGAAAGTGTCTTCAAGAGCTGAATCAA 660
QY 2195 TGGTAGGTTAATGAATTAATCCCACTACTTCA 2231
DB 661 TGGTAGGTTAATGAATTAATCCCACTACTTCA 697

```

RESULT 2

BU741926 676 bp mRNA linear EST 10-OCT-2002
LOCUS BU741926
DEFINITION UI-E-B01-aj-a-j-23-0-UI.s1 UI-E-B01 Homo sapiens cDNA clone
ACCESSION BU741926
VERSION BU741926.1 GI:23687736
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 676)
AUTHORS Bonaldi,M.F., Lennon,G. and Soares,M.B.

TITLE Normalization and subtraction: two approaches to facilitate gene discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE 97044477

COMMENT 8889548

Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9585
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES

source
Location/Qualifiers
1..676
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-B01-aj-a-j-23-0-UI"
/tissue_type="fetal eye"
/dev_stage="fetal"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-B01"
/note="Organ: eye; Vector: pT73-Pec (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-E-B01 is a normalized cDNA library containing the following tissue(s): fetal eye. The library was constructed according to Bonaldi, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pec vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GCGGTATACC. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI).
TAG TISSUE=human fetal eye
TAG LIB=UI-E-B01
TAG_SEQ=GCGGTATACC"

ORIGIN

Query Match 27.0%; Score 657; DB 5; Length 676;
Best Local Similarity 99.1%; Pred. No. 4.6e-132;
Matches 671; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

```

QY 1746 CCTTGAACCATCAGGCGCGTCAAGCTTCAAGTGTACAGGATAGACCCAAAGACA 1805
DB 676 CCTTGAACCATCAGGCGCGTCAAGCTTCAAGTGTACAGGATAGACCCAAAGACA 618
QY 1806 AAAGCTTAAAGACTGTGATTTCTGGGGCCAGAAAGTTTGAATTTCTCTCAGAGAGCG 1865
DB 617 AAAGCTTAAAGACTGTGATTTCTGGGGCCAGAAAGTTTGAATTTCTCTCAGAGAGCG 558
QY 1866 ATGGGCGCTTGTATGAGAGCACCAGAAAGTATTTCAAGTGAAGAGAAACTGTGAGTTTA 1925
DB 557 ATGGGCGCTTGTATGAGAGCACCAGAAAGTATTTCAAGTGAAGAGAAACTGTGAGTTTA 498
QY 1926 ACTGACCGATATGACGAGATATCCCGAAATCACTCTCAAGAACTTGGAAACATCAC 1985
DB 497 ACTGACCGATATGACGAGATATCCCGAAATCACTCTCAAGAACTTGGAAACATCAC 438

```

```

Qy 1986 CAACCAATACACATCACTCAAGATCATGATCTTATGGCTTGGAAAA 2045
Db 437 CAACCAATACACATCACTCAAGATCATGATCTTATGGCTTGGAAAA 378
Qy 2046 GGGATGAAGACATCTCTAAGCTTCTTATGGGATGCTTGGCCAGCGATCTCA 2105
Db 377 GGGATGAAGACATCTCTAAGCTTCTTATGGGATGCTTGGCCAGCGATCTCA 318
Qy 2106 CCTGTGTATACACAGACGACATGATCCAGTCAACGACGATGCTTGGCCAGCTGAA 2165
Db 317 CCTGTGTATACACAGACGACATGATCCAGTCAACGACGATGCTTGGCCAGCTGAA 258
Qy 2166 GAACGTGTCTTCAACACAGCTGAAATCAATGATGTTAATGATTAATAATCCAGACT 2225
Db 257 GAACGTGTCTTCAACACAGCTGAAATCAATGATGTTAATGATTAATAATCCAGACT 198
Qy 2226 ACTTCAAGCTTTAATGCTTTTATTCATTAATACTGTAAGCTGAAACCATTTGA 2285
Db 197 ACTTCAAGCTTTAATGCTTTTATTCATTAATACTGTAAGCTGAAACCATTTGA 138
Qy 2286 AACATTTAATCAACATCTGATTCAGATGCGGAAACCTTATGTTCTATCGAATCCAG 2345
Db 137 AACATTTAATCAACATCTGATTCAGATGCGGAAACCTTATGTTCTATCGAATCCAG 78
Qy 2346 ACAGCCACACCTTAGTATCTGACCAACTAATGATTAATAATAATCAATACCTGTTA 2405
Db 77 ACAGCCACACCTTAGTATCTGACCAACTAATGATTAATAATAATCAATACCTGTTA 18
Qy 2406 AAAAAAAAAAAAAAAAAA 2422
Db 17 AAAAAAAAAAAAAAAAAA 1

RESULT 3
BU730894/c 659 bp mRNA linear EST 09-OCT-2002
LOCUS UI-E-C11-afw-1-11-0-UI.s1 UI-E-C11 Homo sapiens cDNA clone
DEFINITION UI-E-C11-afw-1-11-0-UI 3', mRNA sequence.
ACCESSION BU730894
VERSION BU730894.1 GI:23655244
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 659)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA library preparation: Dr. M. Bento Soares, University of Iowa
cDNA library arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 FORWARD
POLY-A=yes
FEATURES
Source
1..659 Location/Qualifiers
/orf_name="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-C11-afw-1-11-0-UI"

```

```

/tissue_type="RPE and Choroid"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-C11"
/mode="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pT73-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is ACCTA.
This library was created for the program, Gene Discovery
in the Visual System, supported by National Eye Institute
(NEI).
TAG_TISSUE=RPE and Choroid
TAG_LIB=UI-E-C11
TAG_SEQ=ACCTA"

ORIGIN
Query Match 26.8%; Score 651.6; DB 5; Length 659;
Best Local Similarity 99.2%; Pred. No. 6.9e-113;
Matches 654; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1764 CGTCMAAGCTTCACAGGTGTCAAGGATAGACACCAAGACCTTAAGACTGTGA 1823
Db 659 CGTCMAAGCTTCACAGGTGTCAAGGATAGACACCAAGACCTTAAGACTGTGA 600
Qy 1824 GTTCTGGGGCCCAAGAAAGTTTGAATTCCTCAAGAGAGATGGGCTTGAATGAGC 1883
Db 599 GTTCTGGGGCCCAAGAAAGTTTGAATTCCTCAAGAGAGATGGGCTTGAATGAGC 540
Qy 1884 ACCGAGAAGTATCTCAAGTGAAGAGAAACTGTGAGTTTACCTGACGATATGCCAG 1943
Db 539 ACCGAGAAGTATCTCAAGTGAAGAGAAACTGTGAGTTTACCTGACGATATGCCAG 480
Qy 1944 AGATCCCGAAATACCTCAAGAAACCTTGGAAACATCAACCAACATACACACTA 2003
Db 479 AGATCCCGAAATACCTCAAGAAACCTTGGAAACATCAACCAACATACACACTA 420
Qy 2004 CACTTCAAGATCAATGATCTTATTTGGGCTTGGAAAAAGAGATGAACACATTCT 2063
Db 419 CACTTCAAGATCAATGATCTTATTTGGGCTTGGAAAAAGAGATGAACACATTCT 360
Qy 2064 AACCTGCTCTTATTTGGGGAATGCTTGGCAGCCAGGTCCTACCTGTGTACACAGCA 2123
Db 359 AACCTGCTCTTATTTGGGGAATGCTTGGCAGCCAGGTCCTACCTGTGTACACAGCG 300
Qy 2124 GGAACATGATTCAGTCAACGACCATACAGCTGTCCACATGAAAGACGTGCTCAACAG 2183
Db 299 GGAACATGATTCAGTCAACGACCATACAGCTGTCCACATGAAAGACGTGCTCAACAG 240
Qy 2184 CTGAAATCAAAATGTTAGCTTAATAGATTAATAATCCAGACTTACTTACGCTTTAATGCC 2243
Db 239 CTGAAATCAAAATGTTAGCTTAATAGATTAATAATCCAGACTTACTTACGCTTTAATGCC 180
Qy 2244 TTTTATTCATTAATAAAGTGAAGAGTGAAGCACTGGAACATTTGAACATTTAATGAGCTC 2303
Db 179 TTTTATTCATTAATAAAGTGAAGAGTGAAGCACTGGAACATTTGAACATTTAATGAGCTC 120
Qy 2304 TGAATTCAGAGTCCGGGAACCTTAGTTCTATCTGAATTCAGAGACGACACCTTAGTAT 2363
Db 119 TGAATTCAGAGTCCGGGAACCTTAGTTCTATCTGAATTCAGAGACGACACCTTAGTAT 60
Qy 2364 ACTGCCCAACTAATGATTAATTAATAATCAATATCTGTTAATAAAAAAAAAAAAAA 2422
Db 59 ACTGCCCAACTAATGATTAATTAATAATCAATATCTGTTAATAAAAAAAAAAAAAA 1

```

RESULT 4
 BU731809/c
 LOCUS
 DEFINITION BU731809 674 bp mRNA linear EST 09-OCT-2002
 UI-E-C11-alt-g-11-0-UI-s1 UI-E-C11 Homo sapiens cDNA clone
 UI-E-C11-alt-g-11-0-UI 3', mRNA sequence.
 ACCESSION BU731809
 VERSION BU731809.1 GI:23657073
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 674)
 Reference: Bonaldo, M.F., Lennon, G. and Soares, M.B.
 Normalization and subtraction: two approaches to facilitate gene
 discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 PUBMED 8889548
 COMMENT
 Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MBRR, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@uiowa.edu
 Tissue Procurement: Dr. Gregg Hageman
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetica (www.researchgen.com).
 Seq primer: M13 FORWARD
 Polya=yes.

FEATURES

Source Location/Qualifiers
 1..674
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-E-C11-alt-g-11-0-UI"
 /tissue_type="RPE and Choroid"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /clone_1ib="UI-E-C11"
 /note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
 modified polylinker; Site 1: Ecor I; Site 2: Not I;
 UI-E-C11 is a normalized cDNA library containing the
 following tissue(s): RPE and Choroid. The library was
 constructed according to Bonaldo, Lennon and Soares,
 Genome Research, 6:791-806, 1996. First strand cDNA
 synthesis was primed with an oligo-dT primer containing a
 Not I site. Double stranded cDNA was ligated to an Ecor I
 adaptor, digested with Not I, and cloned directionally
 into pT73-Pac vector. The oligonucleotide used to prime
 the synthesis of first-strand cDNA contains a library tag
 sequence that is located between the Not I site and the
 (drr)18 tail. The sequence tag for this library is ACCTA.
 This library was created for the program, Gene Discovery
 in the Visual System, supported by National Eye Institute
 (NEI).
 TAG_TISSUE=RPE and Choroid
 TAG_LIB=UI-E-C11
 TAG_SEQ=ACCTA"

ORIGIN

Query Match 26.7%; Score 648.8; DB 5; Length 674;
 Best Local Similarity 97.9%; Pred. No. 2.8e-130;
 Matches 656; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
 QY 1751 GAACCATCAGCGCGCTCAAGCTTCAAGTGTACAGGACCAACCAAGCAAAAGC 1810
 Db 670 GCACGAGGAGCGCGCTCAAGCTTCAAGTGTACAGGACCAACCAAGCAAAAGC 611

QY 1811 TTAAGACTGTAGTCTGGGGCCAAAGATTGTAATGCTCTCAGAGAGCGATGG 1870
 Db 610 TTAAGACTGTAGTCTGGGGCCAAAGATTGTAATGCTCTCAGAGAGCGATGG 551
 QY 1871 GCGTTGATGAGACACCCAGAAAGTATCTCAAGTGTAGAGAGAAATGTGAGATTAACTG 1390
 Db 550 GCGTTGATGAGACACCCAGAAAGTATCTCAAGTGTAGAGAGAAATGTGAGATTAACTG 491
 QY 1931 ACGATATGCGAGAGATCCCGAAATATCACTCAAAGACCTTTGGAACATCAACACC 1990
 Db 490 ACGATATGCGAGAGATCCCGAAATATCACTCAAAGACCTTTGGAACATCAACACC 431
 QY 1991 AACATACACTTACACTCAAGATCATCATGATGATCTTATGGGCTTGGAAACAGAGAT 2050
 Db 430 AACATACACTTACACTCAAGATCATCATGATGATCTTATGGGCTTGGAAACAGAGAT 371
 QY 2051 GAACGACATTCCTTAACTGCTTCTCTTAATGGGATGCTTGGCCAGCGAGTCTCACTGT 2110
 Db 370 GAACGACATTCCTTAACTGCTTCTCTTAATGGGATGCTTGGCCAGCGAGTCTCACTGT 311
 QY 2111 GTGTACACGAGAGACACTGATCCAGTCAAGCCATACAGCTGTCCACATGAGAACG 2170
 Db 310 GTGTACACGAGAGACACTGATCCAGTCAAGCCATACAGCTGTCCACATGAGAACG 251
 QY 2171 TGTCTTCAACAGCTGATCAATGAATGTTAGCTTAAATGAATAAATCCAGACTACTTC 2230
 Db 250 TGTCTTCAACAGCTGATCAATGAATGTTAGCTTAAATGAATAAATCCAGACTACTTC 191
 QY 2231 AGCCTTAAATGCTTTTAAATCAATAAAGTGAAGAGCTGAGCTGAACTTGAACAT 2290
 Db 190 AGCCTTAAATGCTTTTAAATCAATAAAGTGAAGAGCTGAGCTGAACTTGAACAT 131
 QY 2291 TTAATCTGACTCTGATTCAGAGTGGGAACTTAACTTCAATGATCAACAGACGC 2350
 Db 130 TTAATCTGACTCTGATTCAGAGTGGGAACTTAACTTCAATGATCAACAGACGC 71
 QY 2351 CACACCTGTATGATCGCCCAATCAATGATTTAAATCAATATCTGTTAAATA 2410
 Db 70 CACACCTGTATGATCGCCCAATCAATGATTTAAATCAATATCTGTTAAATA 11
 QY 2411 AAAAAAAAAA 2420
 Db 10 AAAAAAAAAA 1

RESULT 5
 LOCUS BI480798 666 bp mRNA linear EST 28-FEB-2002
 DEFINITION H2RPE-0430 Human Retinal Pigment Epithelium (2) Homo sapiens cDNA
 sequence.
 ACCESSION BI480798
 VERSION BI480798.1 GI:18998607
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 666)
 Reference: Buraczynska, M., Wears, A.V., Zarepari, S., Farjo, R., Filipova, E.,
 Yuan, Y., MacVee, S.P., Hughes, B. and Swaroop, A.
 Towards an expression profile of native human retinal pigment
 epithelium: Identification of a non-redundant set of more than 1100
 genes
 JOURNAL Unpublished (2001)
 COMMENT Contact: Swaroop, A.
 Department of Ophthalmology and Visual Sciences
 Kellogg Eye Center, University of Michigan
 540 KEC, 1000 Wall St., Ann Arbor, MI 48105, USA
 Tel: 734 615 2246
 Fax: 734 647 0228
 Email: swaroop@umich.edu

PCR Primers
FORWARD: M13/PUC-Reverse - ccagcagcagcagctgtcaaacg
BACKWARD: M13/PUC-Forward - agcggatacaacattccacacag
Seq primer: M13/PUC-Reverse.
Location/Qualifiers
1. .666
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="Native Retinal Pigment Epithelium sheets"
/dev_stage="juvenile"
/clone_lib="Human Retinal Pigment Epithelium (2)"
/note="Organ: Retina; Vector: pSport1"

ORIGIN

Query Match 26.4%; Score 640.4; DB 4; Length 666;
Best Local Similarity 99.4%; Pred. No. 1.9e-128;
Matches 652; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
10 CCACCAAGCTAGTGGCCAGACCTTCTGTGGATCATCGGACCCACCTGACCCACCTG 69
12 CCACCAAGCTAGTGGCCAGACCTTCTGTGGATCATCGGACCCACCTGACCCACCTG 71
70 ACCCAAGCCCACTGCTGAGCCCACTGCTGCGCATGACATCACTTACCAAGCCAA 129
72 ACCCAAGCCCACTGCTGAGCCCACTGCTGCGCATGACATCACTTACCAAGCCAA 131
130 TGGCTAATGCGCGCTTGAAGCTCTTCTCCCGCTGCTGCTGCTGCTGCTGCTGCT 189
132 TGGCTAATGCGCGCTTGAAGCTCTTCTCCCGCTGCTGCTGCTGCTGCTGCTGCT 191
190 ACAAGCTGCTATGAGGAGTCTTAAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 249
192 ACAAGCTGCTATGAGGAGTCTTAAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 251
250 TTATAGGCTGCGCCCTCAAGGAGAAACAAGCTGATGTTTGAAGAACTGAATCTGAT 309
252 TTATAGGCTGCGCCCTCAAGGAGAAACAAGCTGATGTTTGAAGAACTGAATCTGAT 311
310 GCGACAGTATCATCAAGTATCCCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 369
312 GCGACAGTATCATCAAGTATCCCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 371
370 TCGGACCCCGCTGAGGAGCAAGTACGAACTGCGCTGCGCCGACCGCTCATGAGCC 429
372 TCGGACCCCGCTGAGGAGCAAGTACGAACTGCGCTGCGCCGACCGCTCATGAGCC 431
430 TGGTGTGCGGCTTCTGCAAGGAGCAAGGAGCAAGGAGCGCTGCTGCGGCAAGCTCA 489
432 TGGTGTGCGGCTTCTGCAAGGAGCAAGGAGCAAGGAGCGCTGCTGCGGCAAGCTCA 491
490 TCGGCTAAGCCCACTGCGGCAAGCTGCTCATCTGCGGAGCTGACGACCGGAGCTTAC 549
492 TCGGCTAAGCCCACTGCGGCAAGCTGCTCATCTGCGGAGCTGACGACCGGAGCTTAC 551
550 AGGCGCTTCCCAAGGAGCAAGTACGAACTGCGGAGCTTATGATCTCGGCAAGCA 609
552 AGGCGCTTCCCAAGGAGCAAGTACGAACTGCGGAGCTTATGATCTCGGCAAGCA 610
610 AGCAGTTGAGAACTGAGCTTACCAACATGTTCTGGGTCGCCCTGAGTGTG 665
611 AGCAGTTGAGAACTGAGCTTACCAACATGTTCTGGGTCGCCCTGAGTGTG 666

RESULT 6
CA389968 639 bp mRNA linear EST 06-NOV-2002
LOCUS CA389968
DEFINITION (Un-normalized, unamplified): cs Homo sapiens cDNA clone csl04h03
5', mRNA sequence.
ACCESSION CA389968
VERSION CA389968.1 GI:24720628
KEYWORDS

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 639)
AUTHORS Wistow,G., Bernstein,S.L., Wyatt,M.K., Farris,R.N., Behal,A.,
Touchman,J.W., Bouffard,G., Smith,D., and Peterson,K.
Expressed sequence tag analysis of human RPE/choroid for the
NEIBank Project: Over 6000 non-redundant transcripts, novel genes
and splice variants
Mol. Vis. 8 (4), 205-220 (2002)
JOURNAL MEDLINE 22103460
PUBMED 12107410
COMMENT Contact: Wistow G
Section on Molecular Structure and Function
National Eye Institute
6/331, NIH, Bethesda, MD 20892-2740, USA
Tel: 301 402 3452
Fax: 301 496 0078
Email: graeme@helix.nih.gov
Plate: 104 row: h column: 03
Seq primer: M13Rpl reverse primer (ABI).
Location/Qualifiers
1. .639
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="csl04h03"
/tissue_type="RPE/choroid"
/dev_stage="Adult"
/lab_host="EMDH10B"
/clone_lib="Human Retinal pigment epithelium/choroid cDNA
(Un-normalized, unamplified): cs"
/note="Organ: Eye; Vector: pCMVSPORTs; Two different donor
eyes (75-80 years old) yielded approximately 600 mg of
dissected RPE/choroid tissue. This in turn yielded 340 ug
of total RNA and 7 ug of mRNA. A directionally cloned cDNA
library in the pCMVSPORTs vector was constructed at Life
Technologies (Rockville, MD: now part of Invitrogen Corp),
essentially following the protocols of the Superscript
Plasmid System (Invitrogen Corp).
<http://www.invitrogen.com/>. The library code
designation was cs. For this library, cDNA inserts were
cloned into the NotI/MluI sites of the vector. EST
analysis was performed on the unamplified library at the
NIH Intramural Sequencing Center (NISC)."

ORIGIN

Query Match 26.3%; Score 638; DB 6; Length 639;
Best Local Similarity 99.8%; Pred. No. 6.3e-128;
Matches 638; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
12 ACCAGCTAGTGGCCAGACCTTCTGTGGATCATGGAACCACTGAAACCCACCTGAC 71
1 ACCAGCTAGTGGCCAGACCTTCTGTGGATCATGGAACCACTGAAACCCACCTGAC 60
72 CCAAGCCCACTGCTGAGCCCACTGCTGCGCATGACATCACTTACCAAGCAAGT 131
61 CCAAGCCCACTGCTGAGCCCACTGCTGCGCATGACATCACTTACCAAGCAAGT 120
132 GCTAATGCGCGCTTGAAGCTCTTCTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCT 191
121 GCTAATGCGCGCTTGAAGCTCTTCTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
192 AAGCTGCTATATGCGGAGTCTTAAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 251
181 AAGCTGCTATATGCGGAGTCTTAAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 240
252 TATAGCTGCGCCCTCAAGGAGAAACAACACTGATGTTTGAAGAACTGACTGTATG 311
241 TATAGCTGCGCCCTCAAGGAGAAACAACACTGATGTTTGAAGAACTGACTGTATG 300
312 GACAGTATATGCAAGTATCCCATTTCTTCTGCTGCTGCTGCTTCTACGTGACGCTG 371

```

Db      301 GACAGCTACATCCAGCTCATCCCAATTCCTTCCTGCTGGGCTTCTAAGAGCGCTGTC
Qy      372 GTGACCCCGCTGTGAACCACTAGAGAACTCCGCTGGCCGACCGCTCATAGAGCTG
Db      361 GTGACCCCGCTGTGAACCACTAGAGAACTCCGCTGGCCGACCGCTCATAGAGCTG
Qy      432 GTGTCGGGCTTCTGTCGAAGCAAGAGAGCAAGCGGCTGCTGGCCGACCGCTCATC
Db      421 GTGTCGGGCTTCTGTCGAAGCAAGAGAGCAAGCGGCTGCTGGCCGACCGCTCATC
Qy      492 CGCTACCCCAACCTGGGCAACGTCTCATCTGCGAGCGCTGACACCGGACGTCTCAAG
Db      481 CGCTACCCCAACCTGGGCAACGTCTCATCTGCGAGCGCTGACACCGGACGTCTCAAG
Qy      552 CGCTTCCCGACCGCCGACGCTGTCTATCTGCGAGCGCTGACACCGGACGTCTCAAG
Db      541 CGCTTCCCGACCGCCGACGCTGTCTATCTGCGAGCGCTGACACCGGACGTCTCAAG
Qy      612 CAGTTGAGAACTAGAGCTTACCAACACATGTTCTGG
Db      601 CAGTTGAGAACTAGAGCTTACCAACACATGTTCTGG

```

RESULT 7
CA397981 636 bp mRNA linear EST 06-NOV-2002
LOCUS CA397981
DEFINITION cs98c05.y1 Human Retinal pigment epithelium/choroid cDNA
(Un-normalized, unamplified): cs Homo sapiens cDNA clone cs98c05
5', mRNA sequence.

ACCESSION CA397981
VERSION CA397981
KEYWORDS EST.
SOURCE GI:24735789

ORGANISM Homo sapiens (human)
Homosapiens

REFERENCE Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 636)

TITLE Wistow G., Bernslein, S.L., Wyatt, M.K., Pariss, R.N., Behal, A.,
Touchman, J.W., Bouffard, G., Smith, D. and Peterson, K.
Expressed sequence tag analysis of human RPE/choroid for the
Neibank Project: Over 6000 non-redundant transcripts, novel genes
and splice variants
Mol. Vis. 8 (4), 205-220 (2002)

JOURNAL 12107410
MEDLINE
PUBMED

COMMENT Contact: Wistow G
Section on Molecular Structure and Function
National Eye Institute
6/331, NIH, Bethesda, MD 20892-2740, USA
Tel: 301 402 3452
Fax: 301 496 0078
Email: gireme@helix.nih.gov

FEATURES
Seq primer: M13RP1 reverse primer (ABI).
Location/Qualifiers
1..636

SOURCE

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="cs98c05"
/tissue_type="RPE/choroid"
/dev_stage="Adult"
/lab_host="EMDH10B"
/clone_id="Human Retinal pigment epithelium/choroid cDNA
(Un-normalized, unamplified): cs
/note="Organ: Eye; Vector: pCMVSPORT6; Two different donor
eyes (75-80 years old) yielded approximately 600 mg of
dissected RPE/choroid tissue. This in turn yielded 340 ug
of total RNA and 7 ug of mRNA. A directionally cloned cDNA
library in the pCMVSPORT6 vector was constructed at Life
Technologies (Rockville, MD; now part of Invitrogen Corp),
essentially following the protocols of the Superscript

ORIGIN

Query Match 26.0%; Score 631.2; DB 6; Length 636;
Best Local Similarity 99.5%; Pred. No. 1.9e-126;
Matches 633; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Plasmid System (Invitrogen Corp).
<http://www.invitrogen.com/> The library code
designation was cs. For this library, cDNA inserts were
cloned into the NotI/Mut sites of the vector. EST
analysis was performed on the unamplified library at the
NIH Intramural Sequencing Center (NISC)."

```

Qy      1328 CCCGAGCCACAGCCCCCTTACACAGCTGCTTCGCCAGTTCCGTCGAGCTCTTATG
Db      1 CCCAGGACACAGCCCCCTTACACAGCTGCTTCGCCAGTTCCGTCGAGCTCTTATG
Qy      1388 GGTCTCACTTCAACATCAGCTGTAACCAAGAGAGATGAGTTCCAGCCCAATCCGAG
Db      61 GGTCTCACTTCAACATCAGCTGTAACCAAGAGAGATGAGTTCCAGCCCAATCCGAG
Qy      1448 GACGAGGAGATGCTACGCTGACATCATTTGCGCTTCTAGAGCTGACATCCATGAT
Db      121 GACGAGGAGATGCTACGCTGACATCATTTGCGCTTCTAGAGCTGACATCCATGAT
Qy      1508 CACCATCTCCCAAGGCAAACTCAAGACCAAACTAATCTGTGCCCCAAGAGGAATCCCTT
Db      181 CACCATCTCCCAAGGCAAACTCAAGACCAAACTAATCTGTGCCCCAAGAGGAATCCCTT
Qy      1568 CTCACAGAGGAGCTGCCCCAAAACACAGAGGACCAACAGAGTTAGGGGCCAGAGAA
Db      241 CTCACAGAGGAGCTGCCCCAAAACACAGAGGACCAACAGAGTTAGGGGCCAGAGAA
Qy      1628 GACAAACAGGCTGTGAAGCTTAAAGCTGTGAAGCTTAAAGCTGTGAAGCTTAAAG
Db      301 GACAAACAGGCTGTGAAGCTTAAAGCTGTGAAGCTTAAAGCTGTGAAGCTTAAAG
Qy      1688 AGGCCAGGCTACTACAGTCCGCCACAGAGGCGCCCTACGCGCCATGTTCTTCCC
Db      361 AGGCCAGGCTACTACAGTCCGCCACAGAGGCGCCCTACGCGCCATGTTCTTCCC
Qy      1748 CTGAAACATCAAGCGCGCTGAAGCTTCAAGTGTCAAGGCTCAAGGCTCAAGG
Db      421 CTGAAACATCAAGCGCGCTGAAGCTTCAAGTGTCAAGGCTCAAGGCTCAAGG
Qy      1808 AGCTTAAAGACTGATGATGTTGGGGCCCAAGAAAGTTTGAATGCTCTCAGAGCGAT
Db      481 AGCTTAAAGACTGATGATGTTGGGGCCCAAGAAAGTTTGAATGCTCTCAGAGCGAT
Qy      1868 GGGGCTTGTATGAGACACCCAGAAATATCTCAAGTGAAGAGAAATCTGTGAGTTAAC
Db      541 GGGGCTTGTATGAGACACCCAGAAATATCTCAAGTGAAGAGAAATCTGTGAGTTAAC
Qy      1928 CTGACGATATGCGCAGAGATCCCGGAATACCTC
Db      601 CTGACGATATGCGCAGAGATCCCGGAATACCTC

```

RESULT 8

BU731149/c 651 bp mRNA linear EST 09-OCT-2002
LOCUS BU731149
DEFINITION UT-E-C11-afg-j-20-0-UT.s1 UT-E-C11 Homo sapiens cDNA clone
BU731149
ACCESSION BU731149
VERSION BU731149.1 GI:23655752
KEYWORDS EST.
SOURCE Homo sapiens (human)
Homosapiens

ORGANISM Homo sapiens (human)
Homosapiens

REFERENCE Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 651)

TITLE Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery

JOURNAL
MEDLINE
PUBMED
COMMENT

Genome Res. 6 (9), 791-806 (1996)
9704447
8889548
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.reagen.com).
Seq primer: M13 FORWARD
POLY-A=yes.

FEATURES
source

Location/Qualifiers
1..651
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-C11-afq-j-20-0-UI"
/tissue_type="RPE and Choroid"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-C11"
/note="Organ: eye; Vector: pTT73-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-E-C11 is a normalized cDNA library containing the
following tissue(s): RPE and Choroid. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pTT73-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is ACCTA.
This library was created for the program, Gene Discovery
in the Visual System, supported by National Eye Institute
(NEI).
TAG_TISSUE=RPE and Choroid
TAG_LIB=UI-E-C11
TAG_SEQ=ACCTA"

ORIGIN

Query Match 25.8%; Score 626.6; DB 5; Length 651;
Best Local Similarity 99.2%; Pred. No. 1.9e-125;
Matches 629; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1593 ACAAGGCGCCAAACAGAACCTTAGGGGCGAGAGACAAACAGGCGCTTGAAGCTTAAG 1652
DB 651 ACAAGGCGCCAAACAGAACCTTAGGGGCGAGAGACAAACAGGCGCTTGAAGCTTAAG 592
QY 1653 CTGTGAGCGCTTCAAGTCTGGCCCACTGTATCAGAGGCCAGGCTACTTACAGTGGCCCAAC 1712
DB 591 CTGTGAGCGCTTCAAGTCTGGCCCACTGTATCAGAGGCCAGGCTACTTACAGTGGCCCAAC 532
QY 1713 AGAGCGCCCTCAGGCCCACTCCAGTCTTCCCTTGAACCAACAGCGCGCTCAAGC 1772
DB 531 AGAGCGCCCTCAGGCCCACTCCAGTCTTCCCTTGAACCAACAGCGCGCTCAAGC 472
QY 1773 TTCAAGTGTCAAGGCGATAGACACCAAGAGCTTAAAGACTGTGAGTTCTGGGG 1832
DB 471 TTCAAGTGTCAAGGCGATAGACACCAAGAGCTTAAAGACTGTGAGTTCTGGGG 412
QY 1833 CCAAGAAAAGTTTGAATTGCTCTCAGAGAGCGATGGGGCTTGTGAGAGACCCAGAG 1892
DB 411 CCAAGAAAAGTTTGAATTGCTCTCAGAGAGCGATGGGGCTTGTGAGAGACCCAGAG 352

QY 1893 TATCTCAAGTAGAGGAGAAAACCTGTGAGTTAACTTGACCGGATATGCCAGATCCCCG 1952
DB 351 TATCTCAAGTAGAGGAGAAAACCTGTGAGTTAACTTGACCGGATATGCCAGATCCCCG 292
QY 1953 AAAATCACCTTAAAGAACCTTTGGAGACATATCAACCAACATACACTACACTCAAG 2012
DB 291 AAAATCACCTTAAAGAACCTTTGGAGACATATCAACCAACATACACTACACTCAAG 232
QY 2013 ATCAATGAGATCCTTATTTGGGCTTTGGAAAACAGGATGAAAGACATTCCTTACCTGCTT 2072
DB 231 ATCAATGAGATCCTTATTTGGGCTTTGGAAAACAGGATGAAAGACATTCCTTACCTGCTT 172
QY 2073 CCTATGGGAGATGCTTGGCCAGCCAGGCTCTTACCTGTGTGTAGACCCAGAGACACTGA 2132
DB 171 CCTATGGGATGCTTGGCCAGCCAGGCTCTTACCTGTGTGTAGACCCAGAGACACTGA 112
QY 2133 TCAGTCAACGACCATATACGCTGTCCACATGAGAGACGTGTCTTACCAACAGCTGATCA 2192
DB 111 TCAGTCAACGACCATATACGCTGTCCACATGAGAGACGTGTCTTACCAACAGCTGATCA 52
QY 2193 AATGTTAGCTTATATGATTAATAATCCAGACTA 2226
DB 51 AATGTTAGCTTATATGATTAATAATCCAGACTA 18

RESULT 9

BM663028/c 592 bp mRNA linear EST 27-FEB-2002
LOCUS
DEFINITION
UI-E-C10-aad-h-10-0-UI.s1 UI-E-C10 Homo sapiens cDNA clone
UI-E-C10-aad-h-10-0-UI 3', mRNA sequence.
BM663028

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS
TITLE
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery

JOURNAL
MEDLINE
PUBMED
COMMENT

Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.reagen.com).
Seq primer: M13 Forward
POLY-A=yes.

FEATURES
source

Location/Qualifiers
1..592
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-C10-aad-h-10-0-UI"
/tissue_type="RPE and Choroid"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-C10"
/note="Organ: eye; Vector: pTT73-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-E-C10 is a cDNA library containing the following
tissue(s): RPE and Choroid. The library was constructed

According to Bonaldo, Lemon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed, with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is ACCCTA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI). TAG_Lib=UI-B-C10 TAG_SEQ=ACCTA"

Qy 1995 TACACACTACACTCAAGATCATGATTCCTATTGGGCGCTTGGAAAAAGGATG 2051
|||||
Db 541 TACACACTACACTCAAGATCATGATTCCTATTGGGCGCTTGGAAAAAGGATG 597

RESULT 11
BM707948 585 bp mRNA linear EST 28-FEB-2002
LOCUS BM707948
DEFINITION UT-E-C11-8-ft-g-11-0-UI r1 UI-E-C11 Homo sapiens cDNA clone
UT-E-C11-8-ft-g-11-0-UI 5', mRNA sequence.

ACCESSION BM707948
VERSION BM707948.1 GI:19021206
KEYWORDS EST.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 585)
AUTHORS Bernaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE 9704447
PubMed 8889548

COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565

Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.reagen.com).

Seq primer: M13 Reverse.

FEATURES
Location/Qualifiers
1..585

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-C11-8-ft-g-11-0-UI"
/tissue_type="RPE and Choroid"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-C11"
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-E-C11 is a normalized cDNA library containing the
following tissue(s): RPE and Choroid. The library was
constructed according to Bernaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pT73-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is ACCTA.
This library was created for the program, Gene Discovery
in the Visual System, supported by National Eye Institute
(NEI)."

ORIGIN
Query Match 23.7%; Score 575; DB 4; Length 585;
Best Local Similarity 98.8%; Pred No. 3.2e-114;
Matches 578; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1759 AGCGCGTCAAGCTTCACAGTGTACAGGATAGACCAAAAGCTTTAAAGAC 1818
|||||
Db 1 AGCGCGTCAAGCTTCACAGTGTACAGGATAGACCAAAAGCTTTAAAGAC 60

Qy 1819 TGTGAGTTCTGGGGCCAAAGAAAGTTTGAATTGCTCTCAGAGAGCGATGGGCGCTTGAT 1878
|||||
Db 61 TGTGAGTTCTGGGGCCAAAGAAAGTTTGAATTGCTCTCAGAGAGTGTATGGGCGCTTGAT 120

Qy 1879 GGAGCACCAGAGATATCTCAAGTGAAGAGAGAAACCTGTGAGTTTAACTGACGATAT 1938
|||||

Db 121 GGAGCACCAGAGATATCTCAAGTGAAGAGAGAAACCTGTGAGTTTAACTGACGATAT 180
|||||

Qy 1939 GCCAGAGATCCCGGAAATATCACTCAAGAACTTTGGAAATCAATCAACCAATCA 1998
|||||

Db 181 GCCAGAGATCCCGGAAATATCACTCAAGAACTTTGGAAATCAATCAACCAATCA 240
|||||

Qy 1999 CACTACACTCAAGATCATGATTCCTATTGGGCGCTTGGAAAAAGGATGAGACACA 2058
|||||

Db 241 CACTACACTCAAGATCATGATTCCTATTGGGCGCTTGGAAAAAGGATGAGACACA 300
|||||

Qy 2059 TTCTTAACCTGCTTCTTAATGAGGATGCTTGGCCAGCGAGTCTCACTGTGTACAC 2118
|||||

Db 301 TTCTTAACCTGCTTCTTAATGAGGATGCTTGGCCAGCGAGTCTCACTGTGTACAC 360
|||||

Qy 2119 CAGCAGACACTGATTCAGTACAGCCATACAGCTGTCCACATGAGAAAGTGTCTTAC 2178
|||||

Db 361 CAGCAGACACTGATTCAGTACAGCCATACAGCTGTCCACATGAGAAAGTGTCTTAC 420
|||||

Qy 2179 AACAGCGTGAATCAAGTGTAGCTTAATAGATAAAATCCAGACTACTTCAGGCTTTA 2238
|||||

Db 421 AACAGCGTGAATCAAGTGTAGCTTAATAGATAAAATCCAGACTACTTCAGGCTTTA 480
|||||

Qy 2239 ATGCGCTTTATTCATTAATAAACTGTGAAGCTAGACTGAACCTTGAACATTTAACTCA 2298
|||||

Db 481 ATGCGCTTTATTCATTAATAAACTGTGAAGCTAGACTGAACCTTGAACATTTAACTCA 540
|||||

Qy 2299 GACTCTGATTCAGAGTCCGGAACCTTTAGTTCTTATCTGATTCGA 2343
|||||

Db 541 GACTCTGATTCAGAGTCCGGAACCTTTAGTTCTTATCTGATTCGA 585
|||||

RESULT 12
BE385296
LOCUS BE385296 593 bp mRNA linear EST 21-JUL-2000
DEFINITION 60127572F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:3618665 5',
mRNA sequence.

ACCESSION BE385296
VERSION BE385296.1 GI:9330661
KEYWORDS EST.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 593)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.

TITLE NIH-MGC
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov

Tissue Procurement: ATCC/DCTD/DRP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)

DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at: image.lnl.gov

Plate: LNCM287 row: 0 column: 18
High quality sequence stop: 593.

FEATURES
Location/Qualifiers
1..593

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3618665"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_20"

ORIGIN
Query Match 23.7%; Score 575; DB 4; Length 585;
Best Local Similarity 98.8%; Pred No. 3.2e-114;
Matches 578; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1759 AGCGCGTCAAGCTTCACAGTGTACAGGATAGACCAAAAGCTTTAAAGAC 1818
|||||
Db 1 AGCGCGTCAAGCTTCACAGTGTACAGGATAGACCAAAAGCTTTAAAGAC 60

/note="Organ: skin; Vector: pOT87; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGG(G). Size selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

Query Match 23.3%; Score 565.2; DB 2; Length 593;
Best Local Similarity 99.3%; Pred. No. 4.3e-112;
Matches 578; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

1471 CATCATTTGGCGCTTCTTCAAGGCTGAGTCCCATGATCCATCCCTCCAGGCAAACTC 1530
1 CATCATTTGGCGCTTCTTCAAGGCTGAGTCCCATGATCCATCCCTCCAGGCAAACTC 60
1531 AAGGACCAAACTCTGTGGCCCAAGAGGGAATCCCTTCCAGAGGCTTGTCCC-AAAA 1589
61 AAGGACCAAACTCTGTGGCCCAAGAGGGAATCCCTTCCAGAGGCTTGTCCC-AAAA 120
1590 ACCAAGAGGAGGCAAAAGAAAGTGAAGGGGCGAGAGAAAGCAAGAGGCTTGAAGCTTA 1649
121 ACCAAGAGGAGGCAAAAGAAAGTGAAGGGGCGAGAGAAAGCAAGAGGCTTGAAGCTTA 180
1650 AAGCTGTGAGCGCTTCAAGTCTGGCCCACTGATCAGAGGCAAGCTTCAAGTGTGCC 1709
181 AAGCTGTGAGCGCTTCAAGTCTGGCCCACTGATCAGAGGCAAGCTTCAAGTGTGCC 240
1710 CACAGAGCGCCCTCAGGCGCCCACTGATCTTCTCCCTTGAAGACCATCAGCGCTTCA 1769
241 CACAGAGCGCCCTCAGGCGCCCACTGATCTTCTCCCTTGAAGACCATCAGCGCTTCA 300
1770 AAGCTTCAAGTGTCAAGGCAATGACCAAGAAAGCTTAAAGCTGAGTGTG 1829
301 AAGCTTCAAGTGTCAAGGCAATGACCAAGAAAGCTTAAAGCTGAGTGTG 360
1830 GGGCCCAAGAAAGTGTGATGCTCTCAAGAGGCGATGGGCGCTTGAAGAGCAAGCCAG 1889
361 GGGCCCAAGAAAGTGTGATGCTCTCAAGAGGCGATGGGCGCTTGAAGAGCAAGCCAG 420
1890 AAGTATCTCAAGTGAAGAGAAAGTGTGAGTGAATTAACCTGAGGATATGCAAGATCC 1949
421 AAGTATCTCAAGTGAAGAGAAAGTGTGAGTGAATTAACCTGAGGATATGCAAGATCC 480
1950 CCGAAATACCTCAAGAAAGCTTGTGAACAATCAACAACAATACATACATCACTCA 2009
481 CCGAAATACCTCAAGAAAGCTTGTGAACAATCAACAACAATACATACATCACTCA 540
2010 AAGATCAATGAGTCTTATTTGGGCTTGAAGAAAGGATG 2051
541 AAGATCAATGAGTCTTATTTGGGCTTGAAGAAAGGATG 582

RESULT 13
BO879880 963 bp mRNA linear EST 16-AUG-2002
LOCUS BO879880
DEFINITION AGENCOURT 8241531 Lupski dorsal root ganglion Homo sapiens cDNA
clone IMAGE:6180559 5', mRNA sequence.
ACCESSION BO879880
VERSION BO879880.1 GI:22271888
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 963)
NIH-MGC <http://mgs.nci.nih.gov/>
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-i@mail.nih.gov

Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
Plate: LLM13563 row: 1 column: 08
High quality sequence stop: 623.

FEATURES

source
Location/Qualifiers
1..963
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6180559"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/clone_id="Lupski dorsal root ganglion"
/note="Vector: pCMV-SPORT6 (Life Technologies); Site 1:
NotI; Site 2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCAGCGCTCCG-3' and
5'-GACTAGTCTAGATCGGAGCGGCGCCCT(15)-3'. Size selected >
1 kb for average insert length 1.7 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."

ORIGIN

Query Match 23.2%; Score 564.6; DB 5; Length 963;
Best Local Similarity 77.3%; Pred. No. 5.8e-112;
Matches 795; Conservative 0; Mismatches 31; Indels 203; Gaps 1;

445 TCAGAGCAAGAGCGAGCAAGGCGGCTGCGGCGACGCTCATCCGCTACGCAACC 504
26 TCAGAGCAAGAGCGAGCAAGGCGGCTGCGGCGACGCTCATCCGCTACGCAACC 85
505 TGGGCAACGTGCTCATCTCTGCGGAGCTGAGCAAGCGAGCTTACAGCGCTTCCCAAG 564
86 TGGGCAACGTGCTCATCTCTGCGGAGCTGAGCAAGCGAGCTTACAGCGCTTCCCAAG 145
565 CCAGACCTGTGTCAGAGGAGCTTATGACTCCGCGAGAAACAAGAGGCTTGAAGAAAC 624
146 CCAGACCTGTGTCAGAGGAGCTTATGACTCCGCGAGAAACAAGAGGCTTGAAGAAAC 205
625 TGAAGCTTACCAACAATGTTCTGAGTGCCTTGGGTGTGTTTCCAACTGTCAATGA 684
206 TGAAGCTTACCAACAATGTTCTGAGTGCCTTGGGTGTGTTTCCAACTGTCAATGA 265
685 AAGCGTGTGTCAGAGTGAATCCGCGAACCTTATCTCTGTCAGAGCTTGTGAAGAGGA 744
266 AAGCGTGTGTCAGAGTGAATCCGCGAACCTTATCTCTGTCAGAGCTTGTGAAGAGGA 325
745 TGAACACTTGGCTACTAGTGTGACACTGTATGCTCAAGACTGATTAATGATCCAC 804
326 TGAACACTTGGCTACTAGTGTGACACTGTATGCTCAAGACTGATTAATGATCCAC 385
805 TGGTGTATACAGAGTGTGACTGTGGCGGTGTACAGCTTCTTCTGATCTGTCTAGTTG 864
386 TGGTGTATACAGAGTGTGACTGTGGCGGTGTACAGCTTCTTCTGATCTGTCTAGTTG 445
865 GCGGCGAGTTTGAACCAAGCAAGGCTTACCTGGCGCAAGACTGAGACTGCTTGTGTC 924
446 GCGGCGAGTTTGAACCAAGCAAGGCTTACCTGGCGCAAGACTGAGACTGCTTGTGTC 505
925 CCGTTCACAGTTCCTGCAAGTTCCTTCTTATGTTGAGTGTGAGTGTGGGCTTCTCA 984
506 CCGTTCACAGTTCCTGCAAGTTCCTTCTTATGTTGAGTGTGAGTGTGGGCTTCTCA 551
985 GGGCCCTCTGCTGAGGAGCAAGGAGGCTCATGGCCAGAGCTGCTTGAAGAGA 1044

```
Db 552 ----- 551
Qy 1045 GGATGCACTGTCCAGAAAGAAAGTCTCAGCGGTAGAAAGACGACGCCCTGTGGCCCA 1104
Db 552 ----- 551
Qy 1105 CACCTGTAATCCAGACTACTCGGAGAGCTGAGAGAGAAATCGCTGATCCGGAGAGC 1164
Db 552 ----- 551
Qy 1165 GGAGCTGTGTGTGCGAGAGCAAGCTCATCAACCCCTTTGGAGAGATGATGATTTTGA 1224
Db 552 ----- 602
Qy 1225 GACCACTGATTTGTCCAGAGAAATTTGCAAGATGTCTCTTTGGCTGTGATAGATGCA 1284
Db 603 GACCAACTGATTTGTCCAGAGAAATTTGCAAGATGTCTCTTTGGCTGTGATAGATGCA 662
Qy 1285 CCAGGACCTGCTCGATGAGAGCCGAGATGATGATTAAGCCGAGCCAGAGCCGCC 1344
Db 663 CCAGGACCTGCTCGATGAGAGCCGAGATGATGATTAAGCCGAGCCAGAGCCGCC 722
Qy 1345 CTACAGAGCTGCTCCGCCAGATTCGTCGAGCCTCTTTATGAGCTCCACCTTCAACAT 1404
Db 723 CTACAGAGCTGCTCCGCCAGATTCGTCGAGCCTCTTTATGAGCTCCACCTTCAACAT 782
Qy 1405 CAGCTGAAACAAGAGAGATGAGATTCAGAGCCCAATAGAGAGAGAGATGCTCA 1464
Db 783 CAGCTGAAACAAGAGAGATGAGATTCAGAGCCCAATAGAGAGAGAGATGCTCA 842
Qy 1465 CGCTGGCAT 1473
Db 843 CTCCGCCCT 851
```

```
RESULT 14
CD518675 881 bp mRNA linear EST 06-JUN-2003
LOCUS AGNCOURT_14375996 NIH_MGC_181 Homo sapiens cDNA clone
DEFINITION IMAGE:30397964 5', mRNA sequence.
ACCESSION CD518675
VERSION CD518675.1 GI:31450393
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 881)
NIH-MGC http://mgs.nci.nih.gov/.
REFERENCE National Institutes of Health, Mammalian Gene Collection (MGC)
AUTHORS Unpublished (1999)
JOURNAL Contact: Daniela S. Gerhard, Ph.D.
COMMENT Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cga@db-remail.nih.gov
Tissue Procurement: Dr. Michael Brownstein
cDNA Library Preparation: Invitrogen Corp
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM482 row: b column: 21
High quality sequence stop: 652.
Location/Qualifiers
1..881
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30397964"
/tissue_type="White Matter"
```

FEATURES

```
/dev_stage="Unknown"
/lab_host="DH10B-Ton A ( T1 and T5 phage resistances)"
/clone_id="NIH_MGC_181"
/notes="Vector: pCMV-Sport6.1; Site 1: NotI; Site 2: EcoRV
(destroyed); Library is oligo-dT primed and directionally
cloned (EcoRV site is destroyed upon cloning). Average
insert size 1.42 kb. Library was constructed by
(Invitrogen). Note: this is a NIH_MGC Library."
```

```
ORIGIN
Query Match 22.8%; Score 553.2; DB 6; Length 881;
Best Local Similarity 76.5%; Pred. No. 1.8e-109;
Matches 817; Conservative 0; Mismatches 43; Indels 208; Gaps 3;

Qy 570 CACCTGGTGAAGAGAGGCTTATGATCCGGCAGAAACAAGAGAGTGGAGAACTGAGC 629
Db 1 CACCTGGTGAAGAGAGGCTTATGATCCGGCAGAAACAAGAGAGTGGAGAACTGAGC 60
Qy 630 CTACCAACAACATGTTCTGGGTGCCCTGGGTGTGTTTGCACCTGTCAATGAGGCG 689
Db 61 CTACCAACAACATGTTCTGGGTGCCCTGGGTGTGTTTGCACCTGTCAATGAGGCG 120
Qy 690 TGGCTTGAAGGTGCAATCCGGGACCTATCTCTCTCCAGAGCTGTGAAAGAGATGAC 749
Db 121 TGGCTTGAAGGTGCAATCCGGGACCTATCTCTCTCCAGAGCTGTGAAAGAGATGAC 180
Qy 750 ACCTTGGTACTCAGTGTGGAACCTGTATGCTTACGATGATGATGATGATGATGATG 809
Db 181 ACCTTGGTACTCAGTGTGGAACCTGTATGCTTACGATGATGATGATGATGATGATG 240
Qy 810 TATACAGAGTGTGACTGTGTGGCGGTGACAGTCTTCTCTGATGATGATGATGATG 869
Db 241 TATACAGAGTGTGACTGTGTGGCGGTGACAGTCTTCTCTGATGATGATGATGATG 300
Qy 870 CAGTTTCTGAAACCAAGCCAGAGGCTTACCTGGGCAATGAGTGAACCTGTGTGCGGTC 929
Db 301 CAGTTTCTGAAACCAAGCCAGAGGCTTACCTGGGCAATGAGTGAACCTGTGTGCGGTC 360
Qy 930 TTCAAGTTCGCGAGTCTTCTTATGTTGGCTGTGAGAGTGGGCTCTCCAGAGGCC 989
Db 361 TTCAAGTTCGCGAGTCTTCTTATGTTGGCTGTGAGAGTGGGCTCTCCAGAGGCC 401
Qy 990 CTGCTGGGCTGAGGAGCATGAGCCAGAGGGGTATGAGCCAGACTGCTTGAAGACGAGATG 1049
Db 402 ----- 401
Qy 1050 CAGTGTCAAGAAAGAAAGTCTCAGGGGTAGAAAGACAGCCAGCGTGTGCGCACACT 1109
Db 402 ----- 401
Qy 1110 GTAATCCAGCTACTCGGAGAGCTGAGGAGAGAAATCGCTTGAACCCGGAGCGGAGG 1169
Db 402 ----- 401
Qy 1170 TTGTGTGGCAGAGCAGCTCATCAACCCCTTTGGAGAGATGATGATTTTGAAGCA 1229
Db 402 ----- 457
Qy 1230 ACTGATTTGTGACAGGAATTTGAGGTGCTCTGTGGCTGTGAGTGAATGACACAGG 1289
Db 458 ACTGATTTGTGACAGGAATTTGAGGTGCTCTGTGGCTGTGAGTGAATGACACAGG 517
Qy 1290 ACCTGGCTCGGATGAGCGCGGACATGATCTGGATTAAGCCCGAGACCAAGCCCCCTTCA 1349
Db 518 ACCTGGCTCGGATGAGCGCGGACATGATCTGGATTAAGCCCGAGACCAAGCCCCCTTCA 577
Qy 1350 CAGCTGCTTCCGCGCAGTTCGTCGAGGCTCTTTATGAGGCTCCACCTTCAATCAAGCC 1409
Db 578 CAGCTGCTTCCGCGCAGTTCGTCGAGGCTCTTTATGAGGCTCCACCTTCAATCAAGCC 637
Qy 1410 TGAACAAAGAGAGATGAGTTCAGGCCCAATACAGAGAGAGAGAGAGAGATGCTCAAGCTG 1469
Db 638 TGAACAAAGAGAGATGAGTTCAGGCCCAATACAGAGAGAGAGAGAGAGATGCTCAAGCTG 697
```

QY 1470 GCATCATTTGGCGCTTCTCTAGGCTGCAATGCCATGATCACCATCTCCAGG--CAA 1527
 DB 698 GCATCATTTGGCGCTTCTCTAGGCTGCAATGCCATGATCACCATCTCCAGGCGCAAC 757
 QY 1528 CTCAGAGCAACCAACTACTGTGTGGCCCAAGAGGAA--TCCCTTCTCCAGAGGCTCTCC 1584
 DB 758 TCCAGAGCAACCAACTACTGTGTGGCCCAAGAGGAAATTCCTTCTCCAGAGGCTCTGC 817
 QY 1585 CAAAAACCAAGCAGCCAAACAGAGCTTAGGGGCCAGGAAGCAA 1632
 DB 818 CCAAAAAACCCCAAGCAGCCAAACCAAGCTTTTAGGGGCCCA 865
 RESULT 15
 LOCUS CO396179
 DEFINITION AGENCOURT_27872871 NIH_MGC_212 Homo sapiens cDNA clone
 IMAGE:30923539 5', mRNA sequence.
 ACCESSION CO396179
 VERSION CO396179.1 GI:49578095
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 756)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: c396bs-remail.nih.gov
 Tissue Procurement: Mary Hendrix
 cDNA Library Preparation: M. Bento Soares, University of Iowa
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LML)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LML at:
 http://image.llnl.gov
 Plate: NDAMI164 row: 1 column: 20
 High quality sequence stop: 571.
 Location/Qualifiers
 1..756
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30923539"
 /tissue_type="Chondrosarcoma Lung Metastasis cell lines"
 /lab_host="DH10B (TI phage resistant)"
 /clone_lib="NIH_MGC_212"
 /note="Organ: Lung; Vector: pYX-Anc; Site 1: EcoR I;
 Site 2: Not I; The library was constructed according
 Bonaldi, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured RNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with oligo-dT
 primer containing a Not I site. Double strand cDNA was
 size selected according to mRNA size fraction, ligated
 with EcoR I adaptor, digested with Not I and then cloned
 directionally into pYX-Anc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is GATAAGGCCA. Tissue was provided by Mary Hendrix."

ORIGIN

Query Match 22.7%; Score 551.6; DB 7; Length 756;
 Best Local Similarity 99.1%; Pred. No. 3.9e-109;
 Matches 554; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 417 CGGCTCATGAGCGCTGTGTGGGCTTGTGCAAGCAAGCAAGCAAGCGCGGCTGTG 476
 DB 8 CGGCTCATGAGCGCTGTGTGGGCTTGTGCAAGCAAGCAAGCAAGCGCGGCTGTG 67

QY 477 CGGCGACGCTCATCCGCTACGCCAACTGAGCAACGCTGCTCATCTGCGAGGCTGAC 536
 DB 68 CGGCGACGCTCATCCGCTACGCCAACTGAGCAACGCTGCTCATCTGCGAGGCTGAC 127
 QY 537 ACCGCACTTCAAGAGCGCTTCCCAAGGCCCAAGACCTGCTGCAAGAGGCTTTATAC 596
 DB 128 ACCGCACTTCAAGAGCGCTTCCCAAGGCCCAAGACCTGCTGCAAGAGGCTTTATAC 187
 QY 597 CCGGCAAGCAACAGCACTTGAAGAACTGAGCTTACCAACAATGTTCTGGGTGCC 656
 DB 188 CCGGCAAGCAACAGCACTTGAAGAACTGAGCTTACCAACAATGTTCTGGGTGCC 247
 QY 657 TGGGTGTGTTTGGCAAAGCTGTCAATGAGCGGTGCTGAGGTGCAATCCGGACCT 716
 DB 248 TGGGTGTGTTTGGCAAAGCTGTCAATGAGCGGTGCTGAGGTGCAATCCGGACCT 307
 QY 717 ATCTGCTCCAGAGCTGCTGCAACGAGATGAACCTTGGTACTCAGTGTGACACTG 776
 DB 308 ATCTGCTCCAGAGCTGCTGCAACGAGATGAACCTTGGTACTCAGTGTGACACTG 367
 QY 777 TATGCTACGACGTGATTAGTATCCACTGATGATACAGAGTGTGACTGTGGCGGTG 836
 DB 368 TATGCTACGACGTGATTAGTATCCACTGATGATACAGAGTGTGACTGTGGCGGTG 427
 QY 837 TACAGCTTCTTCTGACTTGTCTAGTTGAGCGGAGTTTCTGAACCCAGCAAGGCTTAC 896
 DB 428 TACAGCTTCTTCTGACTTGTCTAGTTGAGCGGAGTTTCTGAACCCAGCAAGGCTTAC 487
 QY 897 CTGAGCATAGCTGACCTGTTGTGCGCTTTCACGTTCTGACGTTCTTCTAT 956
 DB 488 CTGAGCATAGCTGACCTGTTGTGCGCTTTCACGTTCTGACGTTCTTCTAT 547
 QY 957 GTTGGCTGCTGAAGGTGG 975
 DB 548 GTTGGCTGCTGAAGGTGG 566

Search completed: March 26, 2005, 13:32:08
 Job time : 7816.3 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2005, 21:57:50 ; Search time 1259.21 seconds
(without alignments)
11419.125 Million cell updates/sec

Title: US-09-622-964A-4
2429
Perfect score: 1 caggagagtcaccacagccta.....aaaaaaaaaaaaaaaaa 2429
Sequence:

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001s:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2429	100.0	2429	2	AAZ21228 Human CG1
2	2399.2	98.8	2404	12	ADQ84435 Human tum
3	2399.2	98.8	2404	13	ADQ83262 Human tum
4	2013	82.9	2229	2	AAZ21227 Human CG1
5	1541.8	63.5	1758	8	ABZ80972 Human bes
6	1244.2	51.2	1263	2	AAV99722 Human adu
7	1008.2	41.5	1238	9	ADA44960 Human pol
8	658.2	27.1	1916	2	AAZ21229 Mouse CG1
9	642.2	26.4	16125	2	AAZ21226 Human CG1
10	642.2	26.4	16650	5	ABA14559 Human ner
11	642.2	26.4	18530	5	ABA14557 Human ner
12	640.6	26.4	18537	5	ABA14558 Human ner
13	587.8	24.2	1717	6	AB189697 Human pol
14	476.4	19.6	7108	5	ABA14556 Human ner
15	431.4	17.8	2112	6	ABO61177 Human ner
16	431.4	17.8	2137	8	ABZ80973 Human bes
17	431	17.7	1530	8	ABZ80997 Human bes
18	350.4	14.4	1326	8	ACA03889 Human bes
19	341.4	14.1	1948	9	ACC59900 Human REM
20	339.4	14.0	1422	8	ABZ80975 Human bes

21	338.2	13.9	2500	11	ADM02614 Human CDN
22	336.8	13.9	2028	8	ABZ80974 Human bes
23	294.6	12.1	305	5	ABA11094 Human ner
24	243.4	10.0	630	13	ADQ53985 Novel can
25	237	9.8	2861	4	AB110793 Drosoph11
26	234.8	9.7	424	2	AAV86976 EST clone
27	219.2	9.0	1592	10	ADC29982 Human nov
28	213	8.8	620	5	ADL45614 Human ova
29	193	7.9	1608	4	AB112609 Drosoph11
30	189.2	7.8	1292	11	ACN89920 Breast ca
31	188.2	7.7	10760	4	AB110792 Drosoph11
32	175	7.2	1119	8	ABZ71972 Human ter
33	175	7.2	1198	6	ABK83853 Human CDN
34	175	7.2	1198	6	ABN95676 Gene #217
35	175	7.2	1234	5	AA67682 DNA encod
36	175	7.2	1279	5	AA85665 DNA encod
37	175	7.2	2183	5	AA891587 DNA encod
38	154.8	6.4	1526	5	AA872796 DNA encod
39	150.4	6.2	3592	4	AB112636 Drosoph11
40	146.8	6.0	1345	4	AB112637 Drosoph11
41	135.2	5.6	539	12	ACH74388 Human gen
42	134.2	5.5	235	12	ACH8088 Human sec
43	126.8	5.2	1350	4	AA76848 Human sec
44	124.4	5.1	2914	12	ADQ24621 Human sof
45	124.4	5.1	2914	12	ADQ25483 Human sof

ALIGNMENTS

RESULT 1	
AAZ21228	
ID	AAZ21228 standard; cDNA; 2429 BP.
XX	
AC	AAZ21228;
DT	22-NOV-1999 (first entry)
XX	
DE	Human CG1CE long form cDNA sequence.
XX	
KW	CG1CE; Best's macular dystrophy; mutation; diagnosis; detection; BMD;
KM	age-related macular dystrophy; se.
OS	
XX	Homo sapiens.
XX	
FT	Key
FT	Location/Qualifiers
FT	105..1412
FT	/*tag= a
FT	/product= "CG1CE short form protein"
XX	
PN	W09943695-A1.
PD	
PD	02-SEP-1999.
XX	
PF	22-FEB-1999; 99WO-US003790.
XX	
PR	25-FEB-1998; 98US-0075941P.
PR	18-DEC-1998; 98US-0112926P.
XX	
PA	(MERI) MERCK & CO INC.
XX	(UYUP-) UNIV UPPSALA.
PI	Petrukhin K, Caskey CT, Metzker M, Madeline C;
XX	
DR	WPI; 1999-540560/45.
DR	P-PSDB; AAY29954.
XX	
PT	Human and mouse polynucleotides encoding CG1CE polypeptides.
XX	
PS	Claim 2; Fig 4; 67pp; English.
XX	
CC	The present sequence represents the human CG1CE cDNA sequence, which when mutated is responsible for Best's macular dystrophy (BMD).

CC Polynucleotides encoding CGICE are useful for diagnosing whether a
CC patient carries a mutation in the CGICE gene. Normal and mutated CGICE
CC proteins are useful for identifying activators and/or inhibitors of these
CC proteins, in order to treat BMD. The CGICE gene offers a simpler and
CC cheaper method of diagnosing BMD without the need for the presence of the
CC patient. The gene may also be useful to discovering the genetic cause of
CC age-related macular dystrophy

SQ Sequence 2429 BP; 614 A; 694 C; 613 G; 508 T; 0 U; 0 Other;

Query Match 100.0%; Score 2429; DB 2; Length 2429;

Best Local Similarity 100.0%; Pident. No. 0;

Matches 2429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGGAGTCCACACAGCCTAGTCGCCAGACCTTCTGGGATCATGGAGCCCACTGGAA 60
Db 1 CAGGAGTCCACACAGCCTAGTCGCCAGACCTTCTGGGATCATGGAGCCCACTGGAA 60
QY 61 CCCCACTGACCCAAAGCCCACTGTCAGCCCACTGGCCATGACCATCACTTAA 120
Db 61 CCCCACTGACCCAAAGCCCACTGTCAGCCCACTGGCCATGACCATCACTTAA 120
QY 121 CAAGCAAGTGGCTAAATGCCCGCTTAGGCTCTTCCCGCCCTGCTGCTGTGCGGG 180
Db 121 CAAGCAAGTGGCTAAATGCCCGCTTAGGCTCTTCCCGCCCTGCTGCTGTGCGGG 180
QY 181 GCAGCATCTAACAGCTGTATATAGCGAGTTCTTATCTTCCGTGCTGTACTACATCA 240
Db 181 GCAGCATCTAACAGCTGTATATAGCGAGTTCTTATCTTCCGTGCTGTACTACATCA 240
QY 241 TCCGCTTATTTATAGGCTGGCCCTCAGAGGAAACAAAGCTGATGTTGAAACTGA 300
Db 241 TCCGCTTATTTATAGGCTGGCCCTCAGAGGAAACAAAGCTGATGTTGAAACTGA 300
QY 301 CTCTGTATTGCGACAGCTACATCAAGCTCATCCCATTTCTTCGTCGCGCTTCTAG 360
Db 301 CTCTGTATTGCGACAGCTACATCAAGCTCATCCCATTTCTTCGTCGCGCTTCTAG 360
QY 361 TGAAGCTGTGCTGACCCGCTGTGTGAAACAGTACAGAACTCTGCGCCAGCCG 420
Db 361 TGAAGCTGTGCTGACCCGCTGTGTGAAACAGTACAGAACTCTGCGCCAGCCG 420
QY 421 TCATGAGCTGTGTGCGGGCTTCTGCGAAGCGAAGAGCGCGGCTGTGCGGC 480
Db 421 TCATGAGCTGTGTGCGGGCTTCTGCGAAGCGAAGAGCGCGGCTGTGCGGC 480
QY 481 GCAAGCTCATCCGCTACGCGCAACCTGGGCAAGCTGTCTATCTGCGAGCGTACGAC 540
Db 481 GCAAGCTCATCCGCTACGCGCAACCTGGGCAAGCTGTCTATCTGCGAGCGTACGAC 540
QY 541 CAGTCTACAGGCTTCCCGACGCGCCAGCACTGTGTGCAAGAGGCTTATGACTCGG 600
Db 541 CAGTCTACAGGCTTCCCGACGCGCCAGCACTGTGTGCAAGAGGCTTATGACTCGG 600
QY 601 CAGAAACACAGCAGTGGAGAACTGAGCTTACCAACAATGTTCTGGGTGCGCTGG 660
Db 601 CAGAAACACAGCAGTGGAGAACTGAGCTTACCAACAATGTTCTGGGTGCGCTGG 660
QY 661 TGTGTTTGGCAACCTGTCAATGAAAGCGTGTGAGAGTCAATCCGGGACCTTATCC 720
Db 661 TGTGTTTGGCAACCTGTCAATGAAAGCGTGTGAGAGTCAATCCGGGACCTTATCC 720
QY 721 TGTCTCAGAGGCTGTGCAAGAGATGAACACTTGGCTACTCAGTGTGCACTGTATG 780
Db 721 TGTCTCAGAGGCTGTGCAAGAGATGAACACTTGGCTACTCAGTGTGCACTGTATG 780
QY 781 CTTACAGCTGATTAATATCCACTGTGTATACACAGGTGTGACTGTGTGCGGTATGA 840
Db 781 CTTACAGCTGATTAATATCCACTGTGTATACACAGGTGTGACTGTGTGCGGTATGA 840
QY 841 GCTTCTCTGACTGTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900
Db 841 GCTTCTCTGACTGT 900

QY 901 GCCATGACGTGACCTCTGTGCGCCGCTTCAAGCTTCTGACAGTCTTCTATGTTG 960
Db 901 GCCATGACGTGACCTCTGTGCGCCGCTTCAAGCTTCTGACAGTCTTCTATGTTG 960
QY 961 GCTGCTGAAGTGGGCTCTTCCAGGCGCTGCTGGAGCATGCGCAGAGGGGTC 1020
Db 961 GCTGCTGAAGTGGGCTCTTCCAGGCGCTGCTGGAGCATGCGCAGAGGGGTC 1020
QY 1021 ATGGCAGACAGCTCTTGAAGAGATGAGATGAGATGAGAAAGAGTCTCACGGGTAG 1080
Db 1021 ATGGCAGACAGCTCTTGAAGAGATGAGATGAGATGAGAAAGAGTCTCACGGGTAG 1080
QY 1081 AAAGCAGCAGGCGTGTGCGCACACTGTAAATCCAGCTACTCGGAGAGCTGAGCAG 1140
Db 1081 AAAGCAGCAGGCGTGTGCGCACACTGTAAATCCAGCTACTCGGAGAGCTGAGCAG 1140
QY 1141 GAGATGCTTGAACCGGAGAGCGGAGGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1200
Db 1141 GAGATGCTTGAACCGGAGAGCGGAGGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1200
QY 1201 TGAAGAGATGATGATATTTTGAACCAACTGATTTGTGACAGGAAATTTGAGGTTG 1260
Db 1201 TGAAGAGATGATGATATTTTGAACCAACTGATTTGTGACAGGAAATTTGAGGTTG 1260
QY 1261 CTTGTTGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320
Db 1261 CTTGTTGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320
QY 1321 GAATTAAGCCCGAGCACAAGCCCTTACAAGCTGCTTCCGCGCCAGTTCCGTGAGGCTC 1380
Db 1321 GAATTAAGCCCGAGCACAAGCCCTTACAAGCTGCTTCCGCGCCAGTTCCGTGAGGCTC 1380
QY 1381 CTTTATGGGCTCACTTCAACATCAAGCTGAAACAAAGAGAGATGAGATTCACGCCAA 1440
Db 1381 CTTTATGGGCTCACTTCAACATCAAGCTGAAACAAAGAGAGATGAGATTCACGCCAA 1440
QY 1441 TCAGAGAGCAGAGAGATGCTCAAGCTGGGATCAATTTGGCGGCTTCTAGGCTGCA 1500
Db 1441 TCAGAGAGCAGAGAGATGCTCAAGCTGGGATCAATTTGGCGGCTTCTAGGCTGCA 1500
QY 1501 CCAATGATCAACATCTTCCAGGCGCAACTCAAGGACCAAACTGATGAGGAGG 1560
Db 1501 CCAATGATCAACATCTTCCAGGCGCAACTCAAGGACCAAACTGATGAGGAGG 1560
QY 1561 ATCCCTTCTCAAGAGGCTGCGCCCAAAACCAAGGACGCAACCAAGCTTAAAGGG 1620
Db 1561 ATCCCTTCTCAAGAGGCTGCGCCCAAAACCAAGGACGCAACCAAGCTTAAAGGG 1620
QY 1621 CCAAGAAACCAAGAGGCTGGAAGCTTAAAGCTGTGAGACCTTCAAGTCTGAGCCACT 1680
Db 1621 CCAAGAAACCAAGAGGCTGGAAGCTTAAAGCTGTGAGACCTTCAAGTCTGAGCCACT 1680
QY 1681 GTATGAGAGGCGGAGCTACTACAGTGTGCGCCCAAGAGCGCCCTTCAAGCTTCA 1740
Db 1681 GTATGAGAGGCGGAGCTACTACAGTGTGCGCCCAAGAGCGCCCTTCAAGCTTCA 1740
QY 1741 CTTCCCTTGAACCACTAGGCGGCTCAAGCTTCAAGTGTCAAGGATTAAGACCAAA 1800
Db 1741 CTTCCCTTGAACCACTAGGCGGCTCAAGCTTCAAGTGTCAAGGATTAAGACCAAA 1800
QY 1801 AGACAAAAGCTTAAAGCTGTGAGTTCTGGGCGCAAGAAAAGTTTGAATTTGCTCAGA 1860
Db 1801 AGACAAAAGCTTAAAGCTGTGAGTTCTGGGCGCAAGAAAAGTTTGAATTTGCTCAGA 1860
QY 1861 GAGGAGTGGGCGCTTGTGAGGACCCAGAAAGTATCTCAAGTGTGAGGAAATCTGTGA 1920
Db 1861 GAGGAGTGGGCGCTTGTGAGGACCCAGAAAGTATCTCAAGTGTGAGGAAATCTGTGA 1920
QY 1921 GTTTAACTGACGATATGCGAGATCCCGGAAATCACTCAAGAACTTTTGAACA 1980
Db 1921 GTTTAACTGACGATATGCGAGATCCCGGAAATCACTCAAGAACTTTTGAACA 1980

QY	1991	TTCCACCAACCAACATACACATACACTCAAAAGATACATGATCCTTATATGCGCCTTGGAA	2040
Db	1991	ATCACCAACCAACATACACATACACTCAAAAGATACATGATCCTTATATGCGCCTTGGAA	2040
QY	2041	AAACAGGATGAAAGACACATTCCTTAACCTGCTTCTTAATGCGGATCCTTCGCGACGACAGT	2100
Db	2041	AAACAGGATGAAAGACACATTCCTTAACCTGCTTCTTAATGCGGATCCTTCGCGACGACAGT	2100
QY	2101	CCTCACCTGTGTGTACACCGAGAGAACCTGATCCAGTACACAGCCATACAGCTGTCCACA	2160
Db	2101	CCTCACCTGTGTGTACACCGAGAGAACCTGATCCAGTACACAGCCATACAGCTGTCCACA	2160
QY	2161	CTGAAGAACGTGTCTCTACAAACGCTGATCAATGCTTATGATATCAAAATCCC	2220
Db	2161	CTGAAGAACGTGTCTCTACAAACGCTGATCAATGCTTATGATATCAAAATCCC	2220
QY	2221	AGACTACTTCAGCCTTTTAATGCTTTTATTCATATAAAACCTGGAAGCTAGACTGAACCA	2280
Db	2221	AGACTACTTCAGCCTTTTAATGCTTTTATTCATATAAAACCTGGAAGCTAGACTGAACCA	2280
QY	2281	TTGGAAAACATTTAATCTGACAGCTGCGGATTCAGAGTCGGGAACCCCTTATGTCATGAT	2340
Db	2281	TTGGAAAACATTTAATCTGACAGCTGCGGATTCAGAGTCGGGAACCCCTTATGTCATGAT	2340
QY	2341	CCAAAGACGCGCAACCTTAGTATACCTGCGCCAAACCTAATGATTTAATAATCAAAATCT	2400
Db	2341	CCAAAGACGCGCAACCTTAGTATACCTGCGCGCCAAACCTAATGATTTAATAATCAAAATCT	2400
QY	2401	CGTTAAAAAATTTTTTTTTTTTTTTTTTTT	2429
Db	2401	CGTTAAAAAATTTTTTTTTTTTTTTTTTTT	2429

CC	XX	RESULT 2
CC	XX	ADQ84435
CC	XX	ID ADQ84435 standard; cDNA; 2404 BP.
CC	XX	AC
CC	XX	ADQ84435;
CC	XX	DT 07-OCT-2004 (first entry)
CC	XX	DE Human tumour-associated antigenic target (TAT) cDNA sequence #1249.
CC	XX	KW human; tumour-associated antigenic target; TAT; cytostatic; gene therapy.
CC	XX	KM cancer; cell proliferative disorder; gene; ss.
CC	XX	OS Homo sapiens.
CC	XX	PN MO2004060270-A2.
CC	XX	PD 22-JUL-2004.
CC	XX	PD 15-OCT-2003; 2003WO-US029126.
CC	XX	PR 18-OCT-2002; 2002US-0418988P.
CC	XX	PA (GETH) GENENTECH INC.
CC	XX	PA (WUTD/) WU T D.
CC	XX	PA (ZHOU/) ZHOU Y.
CC	XX	PI Wu TD, Zhou Y;
CC	XX	DR WPI; 2004-534300/51.
CC	XX	PT New nucleic acid molecule and encoded polypeptide, for diagnosing,
CC	XX	PT preventing or treating cell proliferative disorders such as cancer.
CC	XX	PS Claim 1; SEQ ID NO 1249; 5504bp; English.
CC	XX	The present invention describes an isolated tumour-associated antigenic
CC	XX	target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide
CC	XX	sequences (see SEQ ID No:1 to 4622); (b) the full-length coding region of
CC	XX	(a); (c) the complement of (a) or (b); (d) a sequence that has 80%

CC sequence identity to (a)-(c); (or)e) a sequence that hybridises to (a)-
CC (c). Also described: (1) an expression vector comprising the above
CC nucleic acid; (2) a host cell comprising the above expression vector; (3)
CC a process for producing a polypeptide; (4) an isolated polypeptide
CC comprising: (a) an amino acid sequence encoded by any of the above
CC nucleic acid sequences; (b) an amino acid sequence encoded by the full-
CC length coding region of the above nucleic acid sequences; or (c) a sequence
CC having at least 80% identical to (a) or (b); (5) a chimeric polypeptide
CC comprising the above polypeptide fused to a heterologous polypeptide; (6)
CC an isolated antibody that binds to the above polypeptide; (7) a process
CC for producing the antibody; (8) an isolated oligopeptide that binds to
CC the above polypeptide; (9) a tumour-associated antigenic target (TAT)
CC binding organic molecule that binds to the above polypeptide; (10) a
CC composition of matter comprising the above (chimeric) polypeptide,
CC antibody, oligopeptide or TAT binding organic molecule, in combination
CC with a carrier; (11) an article of manufacture comprising a container and
CC the composition of matter contained within the container; (12) methods of
CC inhibiting the growth of a cell that expresses the above protein, where
CC the growth of the cell is at least in part dependent upon a growth
CC potentiating effect of the above protein; (13) a method of
CC therapeutically treating a mammal having a cancerous tumour comprising
CC cells that express the above protein; (14) a method of determining the
CC presence of a protein in a sample suspected of containing the protein
CC described above; (15) methods of diagnosing the presence of a tumour in a
CC mammal; (16) a method for treating or preventing a cell proliferative
CC disorder associated with increased expression or activity of the above
CC protein; and (17) a method of binding an antibody, oligopeptide or
CC organic molecule to a cell that expresses the protein described above.
CC The TAT sequences have cytosolic activites, and can be used in gene
CC therapy. The composition and methods are useful for diagnosing,
CC preventing or treating cancer. The composition is also used for preparing
CC a medicament for the therapeutic treatment or diagnostic detection of a
CC cell proliferative disorder or cancer. The present sequence represents a
CC human TAT cDNA sequence from the present invention.

Sequence 2404 BP; 590 A; 696 C; 611 G; 507 T; 0 U; 0 Other;

Query Match	98.8%	Score 2399.2;	DB 12;	Length 2404;
Best Local Similarity	99.9%;	Pred. No. 0;		
Matches 2401;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0

OY	1	NAGGAGATGCCAACGAGCCTAGTCGGCAGACCTTCTGTGGGATCATCTGGACCCACCTGGAA	60
Db	1	CAGGAGATGCCAACGAGCCTAGTCGGCAGACCTTCTGTGGGATCATCTGGACCCACCTGGAA	60
OY	61	CCCCACCTGACCCCAAGCCCACTGCTGCGAGCCCACTGCGCAAGAACCATCACTTACA	120
Db	61	CCCCACCTGACCCCAAGCCCACTGCTGCGAGCCCACTGCGCAAGAACCATCACTTACA	120
OY	121	CAAGCAAGTGGCTAATGATCCGCGTTAAGGCTCCTTCTCCGCGCTGCTGTGTGTCGGG	180
Db	121	CAAGCAAGTGGCTAATGATCCGCGTTAAGGCTCCTTCTCCGCGCTGCTGTGTGTCGGG	180
OY	181	GGAGATCTACAAGTGCATATATGGCGAGTTCTTAATCTTCCGCTCTGTACTAATCA	240
Db	181	GGAGATCTACAAGTGCATATATGGCGAGTTCTTAATCTTCCGCTCTGTACTAATCA	240
OY	241	TCCGCTTATTTATATAGGCTGGCCCTCAGCGAAGAACCAACAGCTGATGTTTGAGAACTGA	300
Db	241	TCCGCTTATTTATATAGGCTGGCCCTCAGCGAAGAACCAACAGCTGATGTTTGAGAACTGA	300
OY	301	CTCTGATATTTGGACAGCTTACATCCAGCTCATATCCCATTTCTTCGTGGCGGGCTTACG	360
Db	301	CTCTGATATTTGGACAGCTTACATCCAGCTCATATCCCATTTCTTCGTGGCGGGCTTACG	360
OY	361	TGACGCTGTGCTGAACCCGCTGCTGTAACCAAGTACGAGAACCTGCGGTGACCCGACCGCC	420
Db	361	TGACGCTGTGCTGAACCCGCTGCTGTAACCAAGTACGAGAACCTGCGGTGACCCGACCGCC	420
OY	421	TCATGAGCTGTGTCTGGGCTTCTGTCAAGAGCGAAGACGACGCAAGGCGGGCTGTCTCGGC	480
Db	421	TCATGAGCTGTGTCTGGGCTTCTGTCAAGAGCGAAGACGACGCAAGGCGGGCTGTCTCGGC	480

QY 481 GACGCTCATCCGCTACGCCCAACCTGGGCAACGTGCTCATCTGCGGAGGTGACGACCG 540
 Db 481 GACGCTCATCCGCTACGCCCAACCTGGGCAACGTGCTCATCTGCGGAGGTGACGACCG 540
 QY 541 CAGCTTCAACAGCGCTTCCCGACGCCCGACGACCTGTGTGCAAGAGCGCTTTATGATCTCCG 600
 Db 541 CAGCTTCAACAGCGCTTCCCGACGCCCGACGACCTGTGTGCAAGAGCGCTTTATGATCTCCG 600
 QY 601 CAGAACACAGGACATTTGAGAACTGAGCGCTACACACATGTTCTGGTGGCCCTGCG 660
 Db 601 CAGAACACAGGACATTTGAGAACTGAGCGCTACACACATGTTCTGGTGGCCCTGCG 660
 QY 661 TGTGTTTGGCAACCTGTCAATGAGGCGTGGCTTGGAGGTGCAATCCGGAACCTATCC 720
 Db 661 TGTGTTTGGCAACCTGTCAATGAGGCGTGGCTTGGAGGTGCAATCCGGAACCTATCC 720
 QY 721 TGCTCCAGAGCGCTGTGAGAGATGAACTGTTGGCTACTCATGTGTGACACTGTATG 780
 Db 721 TGCTCCAGAGCGCTGTGAGAGATGAACTGTTGGCTACTCATGTGTGACACTGTATG 780
 QY 781 CTTACGACTGGAATTAATATCCACTGTGTATACACAGGTGTGATGCTGTGGGTGTACA 840
 Db 781 CTTACGACTGGAATTAATATCCACTGTGTATACACAGGTGTGATGCTGTGGGTGTACA 840
 QY 841 GCTTCTTCTGACTTGTGTATGTTGGGCGGAGTTTGTGAACCCAGCCAAAGGCTTACCCTG 900
 Db 841 GCTTCTTCTGACTTGTGTATGTTGGGCGGAGTTTGTGAACCCAGCCAAAGGCTTACCCTG 900
 QY 901 GCCATGAGCTGACCTGTTGTGCGGCTCTCAACCTTCTGAGATTTCTTCTATGTTG 960
 Db 901 GCCATGAGCTGACCTGTTGTGCGGCTCTCTCAACCTTCTGAGATTTCTTCTATGTTG 960
 QY 961 GCTGCTGAAAGTGGGCTCTTCCAGGGCCCTGCTGGCTGAGGAGTGGCCAGAGGGTTC 1020
 Db 961 GCTGCTGAAAGTGGGCTCTTCCAGGGCCCTGCTGGCTGAGGAGTGGCCAGAGGGTTC 1020
 QY 1021 ATGGCAGAGCGCTGTGAGAGAGGATGACATGTGCAAGAAAGAAAGTGTCTACCGGTAG 1080
 Db 1021 ATGGCAGAGCGCTGTGAGAGAGGATGACATGTGCAAGAAAGAAAGTGTCTACCGGTAG 1080
 QY 1081 AAAGCAGCCAGGCGTGTGGGCGCACACCTGTAACTCCAGCTACTCGGAGGCTGAGGAG 1140
 Db 1081 AAAGCAGCCAGGCGTGTGGGCGCACACCTGTAACTCCAGCTACTCGGAGGCTGAGGAG 1140
 QY 1141 GAGATTCGTTGAACCCCGGAGGCGGAGGTTGTGTGCAAGAGCAAGCTCAACCCCTT 1200
 Db 1141 GAGATTCGTTGAACCCCGGAGGCGGAGGTTGTGTGCAAGAGCAAGCTCAACCCCTT 1200
 QY 1201 TGGAGAGATGATGATGATTTTGAAGACCACTGATGTGCGACGAAATTTGCGAGGTTC 1260
 Db 1201 TGGAGAGATGATGATGATTTTGAAGACCACTGATGTGCGACGAAATTTGCGAGGTTC 1260
 QY 1261 CCTGTTGGCTGTGATGATGATGACACAGAGCTTGTGTGAGAGCGGACATGTACTG 1320
 Db 1261 CCTGTTGGCTGTGATGATGATGACACAGAGCTTGTGTGAGAGCGGACATGTACTG 1320
 QY 1321 GAATAAGCCGAGCGACAGCCCGCTTACAGAGCTGCTTCCGCCCAAGTTCCGAGGCTTC 1380
 Db 1321 GAATAAGCCGAGCGACAGCCCGCTTACAGAGCTGCTTCCGCCCAAGTTCCGAGGCTTC 1380
 QY 1381 CTTTATGGGCTCCACTTCAACATCAGCTGTGAACAAAGAGAGATGAGTTCCAGGCCAA 1440
 Db 1381 CTTTATGGGCTCCACTTCAACATCAGCTGTGAACAAAGAGAGATGAGTTCCAGGCCAA 1440
 QY 1441 TCAGAGAGAGAGAGAGATGCTACCGCTGATCATTTGCGCGCTTCTAGAGCTGCAAGTC 1500
 Db 1441 TCAGAGAGAGAGAGAGATGCTACCGCTGATCATTTGCGCGCTTCTAGAGCTGCAAGTC 1500
 QY 1501 CCATGATCACCATCTCTCCAGGGCAAACTCAAGGACCAAACTACTGTGGCCCAAGGGGA 1560
 Db 1501 CCATGATCACCATCTCTCCAGGGCAAACTCAAGGACCAAACTACTGTGGCCCAAGGGGA 1560
 QY 1561 ATCCCTTCTCCAGAGGGGCTGCCCCAAMAAACCAAGGAGCCCAACAGACGTTAGGGG 1620

Db 1561 ATCCCTTCTCCAGAGGGGCTGCCCCAAMAAACCAAGGAGCCCAACAGACGTTAGGGG 1620
 QY 1621 CCAGAGAGACAAAGGCTGTGAAGCTTAAGCTGTGAGGCTTCAAGTCTGGCCACT 1680
 Db 1621 CCAGAGAGACAAAGGCTGTGAAGCTTAAGCTGTGAGGCTTCAAGTCTGGCCACT 1680
 QY 1681 GTATCAGAGGCGAGGCTTACTACAGTGGCCCAAGAGCCCTCGAGCCCACTCCCATGTT 1740
 Db 1681 GTATCAGAGGCGAGGCTTACTACAGTGGCCCAAGAGCCCTCGAGCCCACTCCCATGTT 1740
 QY 1741 CTTCCCTTAACCAATCAGGCGCTCAAAAGCTTCAAGTGTCAAGGCTACAGGCTACACCAA 1800
 Db 1741 CTTCCCTTAACCAATCAGGCGCTCAAAAGCTTCAAGTGTCAAGGCTACAGGCTACACCAA 1800
 QY 1801 AGACAAAGCTTAAAGCTGATGATTTGGGGCCCAAGAAAGTTTGAATGTCTCAAG 1860
 Db 1801 AGACAAAGCTTAAAGCTGATGATTTGGGGCCCAAGAAAGTTTGAATGTCTCTCAAG 1860
 QY 1861 GAGGATGGGGCTTGTATGAGACCCAGAAATATCTCAAGTGAAGAGGAAACCTGTGA 1920
 Db 1861 GAGGATGGGGCTTGTATGAGACCCAGAAATATCTCAAGTGAAGAGGAAACCTGTGA 1920
 QY 1921 GTTTAACCAGATATGCGAGAGATCCCGGAAATACCTCAAGAAACCTTTGGAACA 1980
 Db 1921 GTTTAACCAGATATGCGAGAGATCCCGGAAATACCTCAAGAAACCTTTGGAACA 1980
 QY 1981 ATCAACCAACCAATACACTACTCAAGATCATGATCTTATTTGGGCGCTTGA 2040
 Db 1981 ATCAACCAACCAATACACTACTCAAGATCATGATCTTATTTGGGCGCTTGA 2040
 QY 2041 AAACAGGAGTGAAGCAATTTCTTAACCTGCTTCTAATGGGATGCTTCCAGGCAAGT 2100
 Db 2041 AAACAGGAGTGAAGCAATTTCTTAACCTGCTTCTAATGGGATGCTTCCAGGCAAGT 2100
 QY 2101 CTTACCTGTGTGATACACAGAGGACACTGATTCAGACCATTCAGCTGTCCACA 2160
 Db 2101 CTTACCTGTGTGATACACAGAGGACACTGATTCAGACCATTCAGCTGTCCACA 2160
 QY 2161 CTGAAGAAAGTGTCTTACACAGCGCTGAATCAATGATGTTAGTTAATGAATAAATCCC 2220
 Db 2161 CTGAAGAAAGTGTCTTACACAGCGCTGAATCAATGATGTTAGTTAATGAATAAATCCC 2220
 QY 2221 AGACTACTCAGCTTAAATGCTTTAATGATTAATAAATCTGGAAGCTAAGCTGAACCA 2280
 Db 2221 AGACTACTCAGCTTAAATGCTTTAATGATTAATAAATCTGGAAGCTAAGCTGAACCA 2280
 QY 2281 TTGGAACATTTAATCTAGACTCTGATTCAGAGTCCGGAACCTTATCTGAT 2340
 Db 2281 TTGGAACATTTAATCTAGACTCTGATTCAGAGTCCGGAACCTTATCTGAT 2340
 QY 2341 CCAAGACGCGCACACTTAATGATGATGATGATGATGATGATGATGATGATGATGAT 2400
 Db 2341 CCAAGACGCGCACACTTAATGATGATGATGATGATGATGATGATGATGATGATGAT 2400
 QY 2401 CGTT 2404
 Db 2401 CGTT 2404

RESULT 3
 AD083262 standard; cDNA; 2404 BP.
 AD083262;
 07-OCT-2004 (first entry)
 Human tumour-associated antigenic target (TAT) cDNA sequence #76.
 human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;
 cancer; cell proliferative disorder; gene; ss.

OS Homo sapiens.
XX WO2004060270-A2.
PN 22-JUL-2004.
XX 15-OCT-2003; 2003WO-US029126.
PF 18-OCT-2002; 2002US-0418988P.
PR (GETH) GENENTECH INC.
PA (WUTD/) WU T D.
PA (ZHOU/) ZHOU Y.
XX Mu TD, Zhou Y;
PI WPI; 2004-534300/51.
DR New nucleic acid molecule and encoded polypeptide, for diagnosing,
XX preventing or treating cell proliferative disorders such as cancer.
PT Claim 1; SEQ ID NO 76; 5504bp; English.
XX
XX The present invention describes an isolated tumour-associated antigenic
CC target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide
CC sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of
CC (a); (c) the complement of (a) or (b); (d) a sequence that has 80%
CC sequence identity to (a)-(c); or (e) a sequence that hybridises to (a)-
CC (c). Also described: (1) an expression vector comprising the above
CC nucleic acid; (2) a host cell comprising the above expression vector; (3)
CC a process for producing a polypeptide; (4) an isolated polypeptide
CC comprising: (a) an amino acid sequence encoded by any of the above
CC nucleotide sequences; (b) an amino acid sequence encoded by the full-
CC length coding region of the above nucleotide sequences; or (c) a sequence
CC having at least 80% identical to (a) or (b); (5) a chimeric polypeptide
CC comprising the above polypeptide fused to a heterologous polypeptide; (6)
CC an isolated antibody that binds to the above polypeptide; (7) a process
CC for producing the antibody; (8) an isolated oligopeptide that binds to
CC the above polypeptide; (9) a tumour-associated antigenic target (TAT)
CC binding organic molecule that binds to the above polypeptide; (10) a
CC composition of matter comprising the above (chimeric) polypeptide,
CC antibody, oligopeptide or TAT binding organic molecule, in combination
CC with a carrier; (11) an article of manufacture comprising a container and
CC the composition of matter contained within the container; (12) methods of
CC inhibiting the growth of a cell that expresses the above protein, where
CC the growth of the cell is at least in part dependent upon a growth
CC potentiating effect of the above protein; (13) a method of
CC therapeutically treating a mammal having a cancerous tumour comprising
CC cells that express the above protein; (14) a method of determining the
CC presence of a protein in a sample suspected of containing the protein
CC described above; (15) methods of diagnosing the presence of a tumour in a
CC mammal; (16) a method for treating or preventing a cell proliferative
CC disorder associated with increased expression or activity of the above
CC protein; and (17) a method of binding an antibody, oligopeptide or
CC organic molecule to a cell that expresses the protein described above.
CC The TAT sequences have cytostatic activities, and can be used in gene
CC therapy. The composition and methods are useful for diagnosing,
CC preventing or treating cancer. The composition is also used for preparing
CC a medicament for the therapeutic treatment or diagnostic detection of a
CC cell proliferative disorder or cancer. The present sequence represents a
CC human TAT cDNA sequence from the present invention.
XX
SQ Sequence 2404 BP, 590 A, 696 C, 611 G, 507 T, 0 U, 0 Other;
Query Match 98.8%; Score 2399.2; DB 13; Length 2404;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2401; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CAGGAGTCCCAACGAGCTAGTCCGACAGCTTCTGTGGATCATCGAACCCACTGGAA 60
Db 1 CAGGAGTCCCAACGAGCTAGTCCGACAGCTTCTGTGGATCATCGAACCCACTGGAA 60
QY 61 CCCACCTGACCAAGCCACGCTGTGACGAGCCCACTGCTGCGCATGACATCACTTCA 120

Db 61 CCCACCTGACCAAGCCACGCTGTGACGAGCCCACTGCTGCGCATGACATCACTTCA 120
QY 121 CAGGCAAGTGGCTAAATGCGCGCTTAGGCTCTTCTCCGCGCTGCTGATGCGGG 180
Db 121 CAGGCAAGTGGCTAAATGCGCGCTTAGGCTCTTCTCCGCGCTGCTGATGCGGG 180
QY 181 GCAGCATCTACAACTGCTATATGCGAGTTCCTTAATCTTCTGCTCTGCTACTACATCA 240
Db 181 GCAGCATCTACAACTGCTATATGCGAGTTCCTTAATCTTCTGCTCTGCTACTACATCA 240
QY 241 TCGGCTTTATTTATAGCTGCGCTCAACGGAACAACAGCTATGTTGAAACTGA 300
Db 241 TCGGCTTTATTTATAGCTGCGCTCAACGGAACAACAGCTATGTTGAAACTGA 300
QY 301 CTCTGTATTTGGAAGAGCTATCATCCAGCTCATCCCACTTCTGCTGCTGCTTCAAG 360
Db 301 CTCTGTATTTGGAAGAGCTATCATCCAGCTCATCCCACTTCTGCTGCTGCTTCAAG 360
QY 361 TGAAGCTGTGTGAGACCGCTGTGTGAACAGTACAGAACTGCGGCGGACGCGC 420
Db 361 TGAAGCTGTGTGAGACCGCTGTGTGAACAGTACAGAACTGCGGCGGACGCGC 420
QY 421 TCATGAGCTGTGTGTGCGGCTTGTGTGAAGGCAAGACAGCAAGCCGCTGCTGCGC 480
Db 421 TCATGAGCTGTGTGTGCGGCTTGTGTGAAGGCAAGACAGCAAGCCGCTGCTGCGC 480
QY 481 GCAGGCTCATTCGCTAGCGCAACCTGGGCAACGCTCATCTCTGCGAGCCTCAGACCG 540
Db 481 GCAGGCTCATTCGCTAGCGCAACCTGGGCAACGCTCATCTCTGCGAGCCTCAGACCG 540
QY 541 CAGTCTCAAGCGGCTTCCCGGCGCCAGCACTGTGTGCAAGAGCTTATGATCCCGG 600
Db 541 CAGTCTCAAGCGGCTTCCCGGCGCCAGCACTGTGTGCAAGAGCTTATGATCCCGG 600
QY 601 CAGAACCAAGCACTGTGAGAACTGAGCTTACCAACAACATGTTCTGAGTCCCTGG 660
Db 601 CAGAACCAAGCACTGTGAGAACTGAGCTTACCAACAACATGTTCTGAGTCCCTGG 660
QY 661 TGTGTTTCCCACTGCTGATGAAAGCGCTGTGTGAGAGTCCGGAACCTTATCC 720
Db 661 TGTGTTTCCCACTGCTGATGAAAGCGCTGTGTGAGAGTCCGGAACCTTATCC 720
QY 721 TGTCCAGAGCTGTGTAAGAGATGAACAACCTTGCTGACTCATGTGTGAACCTGTATG 780
Db 721 TGTCCAGAGCTGTGTAAGAGATGAACAACCTTGCTGACTCATGTGTGAACCTGTATG 780
QY 781 CCTACGACTGATTAATGATCCCACTGTGTATACAGAGTGTGAATCTGTGCGGTATCA 840
Db 781 CCTACGACTGATTAATGATCCCACTGTGTATACAGAGTGTGAATCTGTGCGGTATCA 840
QY 841 GCTTCTTCTGATCTTGTCTAATGTTGGGCGGAGTTCGTAACCCAGCCAGGCTTACCTG 900
Db 841 GCTTCTTCTGATCTTGTCTAATGTTGGGCGGAGTTCGTAACCCAGCCAGGCTTACCTG 900
QY 901 GGCATGAGCTGAGCTGTGTGTGCGCGCTTCAAGTTCCTGAGTCTTCTTATATGTTG 960
Db 901 GGCATGAGCTGAGCTGTGTGTGCGCGCTTCAAGTTCCTGAGTCTTCTTATATGTTG 960
QY 961 GCTGCTGTAAGTGGGCTCTCCAGGAGCCGCTGCGGTGAGGCAATGCGAGAGGCTC 1020
Db 961 GCTGCTGTAAGTGGGCTCTCCAGGAGCCGCTGCGGTGAGGCAATGCGAGAGGCTC 1020
QY 1021 ATGCGCAGAGCTGCTTGAAGAGATGAGATGCTCATGAAAGAGAGTCTTCAAGGGTATG 1080
Db 1021 ATGCGCAGAGCTGCTTGAAGAGATGAGATGCTCATGAAAGAGAGTCTTCAAGGGTATG 1080
QY 1081 AAGACAGCGAGCGTGTGTGCGGACACCTGTAATCCAGCTACTCGGAGGCTGAGCGAG 1140
Db 1081 AAGACAGCGAGCGTGTGTGCGGACACCTGTAATCCAGCTACTCGGAGGCTGAGCGAG 1140
QY 1141 GAGATGCTTTGAACCCGAGAGCGGAGTGTGTGTGAGAGAGCTCATCAACCCCTT 1200
Db 1141 GAGATGCTTTGAACCCGAGAGCGGAGTGTGTGTGAGAGAGCTCATCAACCCCTT 1200

```

Db      1141 GAGAAATCGCTTGAAACCCGAGGAGCGAGGTTGTGTGGCAGAGCAGCTCATCAACCCCTT 1200
QY      1201 TGGAGAGATGATGATGATTTTGGAGACCAACTGAGTTGTCCAGAGAAATTTGCGAGTCTC 1260
Db      1201 TGGAGAGATGATGATGATTTTGGAGACCAACTGAGTTGTCCAGAGAAATTTGCGAGTCTC 1260
QY      1261 CCGTGGCTGTGTGATGATGATGACACAGAGACCTGCTCGATGAGAGCGGACATGTAAGT 1320
Db      1261 CCGTGGCTGTGTGATGATGATGACACAGAGACCTGCTCGATGAGAGCGGACATGTAAGT 1320
QY      1321 GAAATAGCCCGAGGCGACAGCCCCCTTACAGAGCTGCTTCCGCCCAATTTCCGTGAGCTTC 1380
Db      1321 GAAATAGCCCGAGGCGACAGCCCCCTTACAGAGCTGCTTCCGCCCAATTTCCGTGAGCTTC 1380
QY      1381 CTTATATGGGCTTCACTTCAACATCACTGATCAAGCAAGAGAGATGAGTTCCAGGCCCA 1440
Db      1381 CTTATATGGGCTTCACTTCAACATCACTGATCAAGCAAGAGAGATGAGTTCCAGGCCCA 1440
QY      1441 TCAGAGAGAGAGAGAGATGCTCAGCTGGCATCATTTGCGCTTCTTAGGCTTGCAGTTC 1500
Db      1441 TCAGAGAGAGAGAGAGATGCTCAGCTGGCATCATTTGCGCTTCTTAGGCTTGCAGTTC 1500
QY      1501 CCAATGATCACCATCTCTCCAGGCGCAACTCAAGAGCCAACTAATCTGTGGCCCAAGAGGA 1560
Db      1501 CCAATGATCACCATCTCTCCAGGCGCAACTCAAGAGCCAACTAATCTGTGGCCCAAGAGGA 1560
QY      1561 ATCCCTTCTCTCAAGAGGCTGCTGCCCAAAAACAAGAGCCCAAGAGCCCAAGAGGGA 1620
Db      1561 ATCCCTTCTCTCAAGAGGCTGCTGCCCAAAAACAAGAGCCCAAGAGCCCAAGAGGGA 1620
QY      1621 CCAGAGAGAGAGAGAGAGGCTTGAAGGCTGTGAGAGCGCTTAAATCTGTGGCCCACT 1680
Db      1621 CCAGAGAGAGAGAGAGAGGCTTGAAGGCTGTGAGAGCGCTTAAATCTGTGGCCCACT 1680
QY      1681 GTATCAGAGGCGCAGGCTACTACTAGTCCCCCAAGAGCGCCCTCAGGCCCACTCCATGTT 1740
Db      1681 GTATCAGAGGCGCAGGCTACTACTAGTCCCCCAAGAGCGCCCTCAGGCCCACTCCATGTT 1740
QY      1741 CTTCCCCCTTGAAGCACTCAAGCCGCTGCAAGGCTTCAAGTGTCAAGGATGAGACCA 1800
Db      1741 CTTCCCCCTTGAAGCACTCAAGCCGCTGCAAGGCTTCAAGTGTCAAGGATGAGACCA 1800
QY      1801 AGACAAAGAGCTTAAAGAGTGTGAGTCTGTGGGCGCAAGAAATTTTGAATGCTCTCAGA 1860
Db      1801 AGACAAAGAGCTTAAAGAGTGTGAGTCTGTGGGCGCAAGAAATTTTGAATGCTCTCAGA 1860
QY      1861 GAGCGATGGGGCTTGTGATGAGAGCACCCAGAAAGTATCTCAAGTGAAGAGAAATCTGTGA 1920
Db      1861 GAGCGATGGGGCTTGTGATGAGAGCACCCAGAAAGTATCTCAAGTGAAGAGAAATCTGTGA 1920
QY      1921 GTTTAACTGACGAGATATGGCCAGAGATCCCGGAAATCACTTATTTGGGCTTGTGA 1980
Db      1921 GTTTAACTGACGAGATATGGCCAGAGATCCCGGAAATCACTTATTTGGGCTTGTGA 1980
QY      1981 ATCACCACCAACATATACACTCACTCAAGATCACTGATCTTATTTGGGCTTGTGA 2040
Db      1981 ATCACCACCAACATATACACTCACTCAAGATCACTGATCTTATTTGGGCTTGTGA 2040
QY      2041 AAAAGGGATGAGAGACATTTCTTAACCTGCTTCTTAATGGGGATGCTTCCGACGAGCT 2100
Db      2041 AAAAGGGATGAGAGACATTTCTTAACCTGCTTCTTAATGGGGATGCTTCCGACGAGCT 2100
QY      2101 CCTCACTGTGTGATGACACAGAGACATGATCCAGTCAAGCAATACAGCTGTCTCACA 2160
Db      2101 CCTCACTGTGTGATGACACAGAGACATGATCCAGTCAAGCAATACAGCTGTCTCACA 2160
QY      2161 CTGAAGAAAGTGTTCTTCAACAGCTGATCAATCAATGTTAGTTAATAGTAAATATCCC 2220
Db      2161 CTGAAGAAAGTGTTCTTCAACAGCTGATCAATCAATGTTAGTTAATAGTAAATATCCC 2220
QY      2221 AGACTACTGAGCTTTAATGCTTTATTCATTAATAAAGCTAAGACTGAACA 2280
Db      2221 AGACTACTGAGCTTTAATGCTTTATTCATTAATAAAGCTAAGACTGAACA 2280

```

```

QY      2281 TTGAAACATTTTAACTGAGACTGTGATTCAGAGTGGGAAACCTTAGTTCTATCTGAT 2340
Db      2281 TTGAAACATTTTAACTGAGACTGTGATTCAGAGTGGGAAACCTTAGTTCTATCTGAT 2340
QY      2341 CCAAGACGACCACTTAGTATGATCTGCCAACTAATGATTTAATTAATAAATACT 2400
Db      2341 CCAAGACGACCACTTAGTATGATCTGCCAACTAATGATTTAATTAATAAATACT 2400
QY      2401 CGTT 2404
Db      2401 CGTT 2404

RESULT 4
AAZ1227
ID  AAZ1227 standard; cDNA; 2229 BP.
XX
XX  AAZ1227;
AC
XX  22-NOV-1999 (first entry)
DT
XX
XX  Human CGICE short form cDNA sequence.
DE
XX
XX  CGICE; Best's macular dystrophy; mutation; diagnosis; detection; BMD;
KW  age-related macular dystrophy; ss.
XX
XX  Homo sapiens.
OS
XX
XX  Key      Location/Qualifiers
FH      CDS
FT      105..1862
FT      /*tag= a
FT      /product= "CGICE long form protein"
FT      /transl_except= (pos:465..467,aa:Ser)

W09943695-A1.
XX
XX  02-SEP-1999.
PD
XX
XX  22-FEB-1999; 99MO-US003790.
PF
XX
XX  25-FEB-1998; 98US-0075941P.
PR  18-DEC-1998; 98US-0112926P.
XX
XX  (MERI ) MERCK & CO INC.
PA  (UYUP-) UNIV UPPSALA.
XX
XX  Petrukhin K, Caskey CT, Metzker M, Wadelius C;
PI
XX
XX  WPI; 1999-540560/45.
DR  P-PSDB; AAY29953.
XX
XX  Human and mouse polynucleotides encoding CGICE polypeptides.
PT
XX
XX  Claim 2; Fig 2; 67pp; English.
PS
XX
XX  The present sequence represents the human CGICE cDNA sequence, which when
CC  mutated is responsible for Best's macular dystrophy (BMD).
CC  Polynucleotides encoding CGICE are useful for diagnosing whether a
CC  patient carries a mutation in the CGICE gene. Normal and mutated CGICE
CC  proteins are useful for identifying activators and/or inhibitors of these
CC  proteins, in order to treat BMD. The CGICE gene offers a simpler and
CC  cheaper method of diagnosing BMD without the need for the presence of the
CC  patient. The gene may also be useful to discovering the genetic cause of
CC  age-related macular dystrophy
XX
XX  Sequence 2229 BP, 575 A; 646 C; 532 G; 476 T; 0 U; 0 Other;

Query Match      82.9%; Score 2013; DB 2; Length 2229;
Best Local Similarity 91.6%; Pred. No. 0;
Matches 2226; Conservative 0; Mismatches 0; Indels 203; Gaps 1;

QY      1 CAGGAGTCCACACAGCTGATGCGCAGACCTTGTGGATCATCGGACCACTTGAA 60

```

|||||
Db 1 CAGGAGTCCACCGAGCTAGTCGCCAGACCCTTGTGGGATCATCGAACCCACCTGGAA 60
Qy 61 CCCACCGACCCAGACCCCACTGCTGAGACCCCACTGCGCCATGACATCACTTACA 120
Db 61 CCCACCGACCCAGACCCCACTGCTGAGACCCCACTGCGCCATGACATCACTTACA 120
Qy 121 CAAGCAAGTGGCTAATGCCGCTTAGGCTCCTTCTCCGCGCTGCTGCTGCTGGCGGG 180
Db 121 CAAGCAAGTGGCTAATGCCGCTTAGGCTCCTTCTCCGCGCTGCTGCTGCTGGCGGG 180
Qy 181 GCAGATCTTACAGCTGCTAATGCGAGATTCTTAATCTTCTGCTGCTAATCAATCA 240
Db 181 GCAGATCTTACAGCTGCTAATGCGAGATTCTTAATCTTCTGCTGCTAATCAATCA 240
Qy 241 TCCGCTTATTTAATAGCTGGGCCCTCAGGAGAACAAACAGCTGATGTTGAGAACTGA 300
Db 241 TCCGCTTATTTAATAGCTGGGCCCTCAGGAGAACAAACAGCTGATGTTGAGAACTGA 300
Qy 301 CTCTGTATTTGCGACAGCTACATCCAGCTCATCCCATTTCTTCTGCTGGGCTTCAAG 360
Db 301 CTCTGTATTTGCGACAGCTACATCCAGCTCATCCCATTTCTTCTGCTGGGCTTCAAG 360
Qy 361 TGACGCTGGTGTGACCCGCTGGTGAACAGTACAGAACCTGCGTGGCCGACCGCC 420
Db 361 TGACGCTGGTGTGACCCGCTGGTGAACAGTACAGAACCTGCGTGGCCGACCGCC 420
Qy 421 TCATGAGCTGGTGTGCGGGCTTCTGCGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
Db 421 TCATGAGCTGGTGTGCGGGCTTCTGCGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
Qy 481 GCACGCTCATCGCTACGACCAACCTGGGCAAGCTCATCTGCGCAAGCTCAGACAG 540
Db 481 GCACGCTCATCGCTACGACCAACCTGGGCAAGCTCATCTGCGCAAGCTCAGACAG 540
Qy 541 CAGTCTACAGCGCTTCCCAAGCGCCCAACCTGCTGCAAGCAGAGCTTATGACTCCGG 600
Db 541 CAGTCTACAGCGCTTCCCAAGCGCCCAACCTGCTGCAAGCAGAGCTTATGACTCCGG 600
Qy 601 CAGAACACACAGCTGGAGAGAACTGAGCTTACCAACATGTTCTGGGCTGCTGG 660
Db 601 CAGAACACACAGCTGGAGAGAACTGAGCTTACCAACATGTTCTGGGCTGCTGG 660
Qy 661 TGTGGTTTGCCACCTGTCAATGAAAGGCGTGGAGTGAATCCGGGACCTATCC 720
Db 661 TGTGGTTTGCCACCTGTCAATGAAAGGCGTGGAGTGAATCCGGGACCTATCC 720
Qy 721 TGCTCCAGAGCGCTGCTGAACGAGATGAACACCTTGCTGCTCACTGAGTGACACTGTATG 780
Db 721 TGCTCCAGAGCGCTGCTGAACGAGATGAACACCTTGCTGCTCACTGAGTGACACTGTATG 780
Qy 781 CCTACGACTGGATTAATATCCCATGCTGTATACACAGCTGGTGACTGTGGCGGTGACA 840
Db 781 CCTACGACTGGATTAATATCCCATGCTGTATACACAGCTGGTGACTGTGGCGGTGACA 840
Qy 841 GCTTCTTCTCACTTGTCTAGTTGGGCGGAGTTCTGAACCCAGCCAGCTACCTCG 900
Db 841 GCTTCTTCTCACTTGTCTAGTTGGGCGGAGTTCTGAACCCAGCCAGCTACCTCG 900
Qy 901 GCGATGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Db 901 GCGATGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Qy 961 GCTGCTGAGAGTGGGCTCTCCAGAGGCGCTGCTGGGCTGAGAGCATACCCAGAGGGGTC 1020
Db 961 GCTGCTGAGAGTGGGCTCTCCAGAGGCGCTGCTGGGCTGAGAGCATACCCAGAGGGGTC 1020
Qy 971 GCTGCTGAGAGTGGGCTCTCCAGAGGCGCTGCTGGGCTGAGAGCATACCCAGAGGGGTC 970
Db 971 GCTGCTGAGAGTGGGCTCTCCAGAGGCGCTGCTGGGCTGAGAGCATACCCAGAGGGGTC 970
Qy 1081 AATGCGACAGCTGCTTGAAGAGAGATGAGTGTCAAGAAAGAAAGTCTCAGCGGTAG 1080
Db 1081 AATGCGACAGCTGCTTGAAGAGAGATGAGTGTCAAGAAAGAAAGTCTCAGCGGTAG 1080
Qy 971 ----- 970
Db 971 ----- 970
Qy 1081 AATGCGACAGCTGCTTGAAGAGAGATGAGTGTCAAGAAAGAAAGTCTCAGCGGTAG 1080
Db 1081 AATGCGACAGCTGCTTGAAGAGAGATGAGTGTCAAGAAAGAAAGTCTCAGCGGTAG 1080

|||||
Db 971 ----- 970
Qy 1141 GAGAAATGCTTGAACCCGGAGGCGAGGTTGTGGTGACAGAGCAGCTCATCAACCCCTT 1200
Db 971 -----GCTGGCAGAGCAGCTCATCAACCCCTT 997
Qy 1201 TGAAGAGATGATGATGATTTTGAAGACCACTGATTTGCAACAGAAATTTGCAAGTGTG 1260
Db 998 TGAAGAGATGATGATGATTTTGAAGACCACTGATTTGCAACAGAAATTTGCAAGTGTG 1057
Qy 1261 CCGTGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320
Db 1058 CCGTGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1117
Qy 1321 GAATPAGCCGAGGACAGCCGCCCTTACACAGCTGCTTCCGCCGATTCGGTCAGGCTTC 1380
Db 1118 GAATPAGCCGAGGACAGCCGCCCTTACACAGCTGCTTCCGCCGATTCGGTCAGGCTTC 1177
Qy 1381 CTTATGCGCTCCACTTCAACATCAGCTGTAACAAAGAGAGATGAGATTCCAGCCCA 1440
Db 1178 CTTATGCGCTCCACTTCAACATCAGCTGTAACAAAGAGAGATGAGATTCCAGCCCA 1237
Qy 1441 TGAAGAGAGAGAGAGATGCTCAGCTGCGATCATTTGGCGCTTCTTAAGCTGCAATG 1500
Db 1238 TGAAGAGAGAGAGAGATGCTCAGCTGCGATCATTTGGCGCTTCTTAAGCTGCAATG 1297
Qy 1501 CCATGATACCATCTCCAGAGGCAAACTCAAGAGACCAACTGATGAGGAGAGAGAGAGAG 1560
Db 1298 CCATGATACCATCTCCAGAGGCAAACTCAAGAGACCAACTGATGAGGAGAGAGAGAGAG 1357
Qy 1561 ATCCCTTCTCCAGAGGCGCTGCCCCAAGAACCAACAGGACAGCCAAACGTTAGGGG 1620
Db 1358 ATCCCTTCTCCAGAGGCGCTGCCCCAAGAACCAACAGGACAGCCAAACGTTAGGGG 1417
Qy 1621 CCAGAGAGACCAAGAGGCTGGAAGCTTGAAGCTGAGAGAGAGAGAGAGAGAGAGAGAG 1680
Db 1418 CCAGAGAGACCAAGAGGCTGGAAGCTTGAAGCTGAGAGAGAGAGAGAGAGAGAGAGAG 1477
Qy 1681 GTATCAGAGGCGAGCTGATCAGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
Db 1478 GTATCAGAGGCGAGCTGATCAGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1537
Qy 1741 CTTCCCTTGAAGACATCAGGCGCTCAAGCTTCAACAGTGTCAAGGACATAGACACCA 1800
Db 1538 CTTCCCTTGAAGACATCAGGCGCTCAAGCTTCAACAGTGTCAAGGACATAGACACCA 1597
Qy 1801 AGACAAAAGCTTAAAGACTGTGAGTCTGGGCGCAAGAAAGTTTGAATGCTCTCAGA 1860
Db 1598 AGACAAAAGCTTAAAGACTGTGAGTCTGGGCGCAAGAAAGTTTGAATGCTCTCAGA 1657
Qy 1861 GAGGAGTGGGCGCTTGAATGAGACCCAGAAAGTATCTCAAGTGAAGAGAGAGAGAGAG 1920
Db 1658 GAGGAGTGGGCGCTTGAATGAGACCCAGAAAGTATCTCAAGTGAAGAGAGAGAGAGAG 1717
Qy 1921 GTTTAACTGACGAGTATGACAGAGATCCCGGAAATCAGCTCAAGAAAGCTTTGGAACA 1980
Db 1718 GTTTAACTGACGAGTATGACAGAGATCCCGGAAATCAGCTCAAGAAAGCTTTGGAACA 1777
Qy 1981 ATCAACCAACCAATACATCACTCAAGATCAATGATCTTAATTTGGGCGCTTGA 2040
Db 1778 ATCAACCAACCAATACATCACTCAAGATCAATGATCTTAATTTGGGCGCTTGA 1837
Qy 2041 AAACAGGAGTGAACACATTTCTTAACCTGCTTCTTAATGAGGAGTGTGCGCAGCAGGT 2100
Db 1838 AAACAGGAGTGAACACATTTCTTAACCTGCTTCTTAATGAGGAGTGTGCGCAGCAGGT 1897
Qy 2101 CCTCAGCTGTGTGACACAGCAGAGACATGATCAAGTCAAGCAGATCAAGCTGTGACA 2160
Db 1898 CCTCAGCTGTGTGACACAGCAGAGACATGATCAAGTCAAGCAGATCAAGCTGTGACA 1957
Qy 2161 CTGAAGAAAGTGTCTTCAACAGCAGCTGAATGAATGTTAGCTTAAATGAATAAATCC 2220
Db 1958 CTGAAGAAAGTGTCTTCAACAGCAGCTGAATGAATGTTAGCTTAAATGAATAAATCC 2017

QY	1065	AAGGCTCACGGGTAGAAAGCACCAGGCGTGTGCGCACACCTGTATATCCAGACTACT	112.2
Db	867	-----	866
QY	1125	CGGAGGCTGAGGCAGAGAAATCGCTTGAACCCGAGAGCGAGGTGTGTGTGCAGAGC	118.4
Db	867	-----	867
QY	1185	AGCTCATCAACCCCTTTTGGAGAGATGTATGATTTTGAAGCCAACTGTATTTGTGCACA	124.4
Db	878	AGCTCATCAACCCCTTTTGGAGAGATGTATGATTTTGAAGCCAACTGTATTTGTGCACA	93.7
QY	1245	GGAAATTGSCAGAGTGCCTGTGTGGCTGTGATGTGATGTCACACAGGACCTGACCTGGATGG	130.4
Db	938	GGAAATTGSCAGAGTGCCTGTGTGGCTGTGATGTGATGTCACACAGGACCTGACCTGGATGG	99.7
QY	1305	AGCCGGAATGTACTGTGAATTAAGCCCGAGCACAAGCCCCCTTACACAGCTGTCTTCGCGCC	136.4
Db	998	AGCCGGAATGTACTGTGAATTAAGCCCGAGCACAAGCCCCCTTACACAGCTGTCTTCGCGCC	105.7
QY	1365	AGTTCCGTGAGGCTCTCTTTATGGGCTTCACCTTCAATCAAGCTTGAACTGAAGAGAGA	142.2
Db	1058	AGTTCCGTGAGGCTCTCTTTATGGGCTTCACCTTCAATCAAGCTTGAACTGAAGAGAGA	111.7
QY	1425	TGAGATTTCAGCCCCAATCAGGAGGACGAGAGGATGCTCAGGCTGGACATCATTTGGCGGCT	148.4
Db	1118	TGAGATTTCAGCCCCAATCAGGAGGACGAGAGGATGCTCAGGCTGGACATCATTTGGCGGCT	117.7
QY	1485	TCCTAGGCTTCAGTCCCGCATGATCAACAATCTCTCCAGGGCAAACTCAAGACAACAACTTAC	154.4
Db	1178	TCCTAGGCTTCAGTCCCGCATGATCAACAATCTCTCCAGGGCAAACTCAAGACAACAACTTAC	123.7
QY	1545	TGTGTGCCCCAAGAGGAATCCCTTCTTCCACAGGGCTGTGCCAAAAACACACAAAGCAGCCA	160.4
Db	1238	TGTGTGCCCCAAGAGGAATCCCTTCTTCCACAGGGCTGTGCCAAAAACACACAAAGCAGCCA	129.7
QY	1605	AACAGAACTGTAGGGGCGCAGGAGAGCAACAAGGCGCTGGAACCTTAAAGGCTGTGAGCGGCT	166.4
Db	1298	AACAGAACTGTAGGGGCGCAGGAGAGCAACAAGGCGCTGGAACCTTAAAGGCTGTGAGCGGCT	135.7
QY	1665	TCAAAGTCTGCGCCACTGTATACAGAGCCAGGCTACTACAGTGTCCCCACAGACGCCCCCTCA	172.4
Db	1358	TCAAAGTCTGCGCCACTGTATACAGAGCCAGGCTACTACAGTGTCCCCACAGACGCCCCCTCA	141.7
QY	1725	GCCCCACGCCATGTCTTCCGCCCTAGAACCATACAGGCGCGTCAAGTCTTCAAGTGTCA	178.4
Db	1418	GCCCCACGCCATGTCTTCCGCCCTAGAACCATACAGGCGCGTCAAGTCTTCAAGTGTCA	147.7
QY	1785	CAGGCAATGACACACAAAGACAAAAGCTTAAAGACTGTAGTTCTGGGCGCCAAAGAAAAGTT	184.4
Db	1478	CAGGCAATGACACACAAAGACAAAAGCTTAAAGACTGTAGTTCTGGGCGCCAAAGAAAAGTT	153.7
QY	1845	TTGAATGTGCTCTCAGAGAGCGATGGGGCTTTGATGTGAGACCCACGAAAGTATCTCAAGTGA	190.4
Db	1538	TTGAATGTGCTCTCAGAGAGCGATGGGGCTTTGATGTGAGACCCACGAAAGTATCTCAAGTGA	159.7
QY	1905	GGAGGAAAACCTGTGGAGTTTAACCTGACGGAATATGCCAGAAATCTCCGAAAATCACTTCA	196.4
Db	1598	GGAGGAAAACCTGTGGAGTTTAACCTGACGGAATATGCCAGAAATCTCCGAAAATCACTTCA	165.7
QY	1965	AAGAACTTTTGGAAACAATCACCAACCAACATACACACTTACACTGAAGATCACATGTATC	202.4
Db	1658	AAGAACTTTTGGAAACAATCACCAACCAACATACACACTTACACTGAAGATCACATGTATC	171.7
QY	2025	CTTATATGGGCTTGTGAAAACAGGATGTAAGCAATTTCTTAA	206.5
Db	1718	CTTATATGGGCTTGTGAAAACAGGATGTAAGCAATTTCTTAA	175.8

RESULT 6	
AAV99722	
ID	AAV99722 standard; cDNA; 1263 BP.

XX	AC	AAV99722;
XX	DT	26-APR-1999 (First entry)
XX	DE	Human adult retina secreted protein bk112_15 cDNA.
XX	KM	Secreted protein; human; retina; bk112_15; ds.
XX	OS	Homo sapiens.
XX	FM	Key
XX	FT	Location/Qualifiers
XX	CD	100..885
XX	FT	/*tag= a
XX	PN	MO9856909-A2.
XX	PD	17-DEC-1998.
XX	PF	08-JUN-1998; 98MO-US011822.
XX	PR	11-JUN-1997; 97US-00873218.
XX	PR	05-JUN-1998; 98US-00092722.
XX	PA	(GENY) GENETICS INST INC.
XX	PI	Jacobs K, McCoy JM, Lavallie EF, Racie LA, Treacy M, Spaulding V;
XX	PI	Agostino MJ, Howes SH, Fecthel K;
XX	DR	WPI; 1999-080899/07.
XX	DR	P-PSDB; AAW95345.
XX	PT	New polynucleotides encoding secreted human proteins - derived from human
XX	PT	foetal brain, adult testes, foetal kidney, adult thyroid or adult retina
XX	PT	cDNA libraries.
XX	PS	Claim 12; Page 71-72; 113pp; English.
XX	CC	This is the nucleotide sequence of cDNA clone bk112_15, which includes an
XX	CC	open reading frame for a 261-amino acid polypeptide (see AAW95345). The
XX	CC	clone was isolated from a human adult retina cDNA library using methods
XX	CC	which are selective for cDNAs encoding secreted proteins, or was
XX	CC	identified as encoding a secreted or transmembrane protein on the basis
XX	CC	of computer analysis of the amino acid sequence of the encoding protein.
XX	CC	Database searches indicate some sequence similarity to known sequences.
XX	CC	The invention provides cDNA clones (see AAV9721-33) from human adult
XX	CC	thyroid, adult retina, adult testis, foetal kidney and foetal brain that
XX	CC	encode novel secreted proteins (see AAW95344-53). Each clone is
XX	CC	individually available from deposit clone ATCC 98451 (see also AAV9724-
XX	CC	43). The isolated polynucleotides (PNS) and proteins are predicted to
XX	CC	have activities which would make them suitable for treating, preventing
XX	CC	or ameliorating medical conditions in humans and animals, although no
XX	CC	supporting data is given. Suggested activities include nutritional, (e.g.
XX	CC	cytokine, cell proliferation or differentiation, immune stimulating, (e.g.
XX	CC	as vaccines) or immune suppressing, haematopoiesis regulating, tissue
XX	CC	growth, activin/inhibin, chemotactic/chemokinetic, haemostatic,
XX	CC	thrombolytic, receptor/ligand, antiinflammatory, cadherin/tumour invasion
XX	CC	suppressor, and tumour inhibition activities. The PNS are also stated to
XX	CC	be useful for gene therapy
XX	XX	Sequence 1263 BP; 397 A; 342 C; 272 G; 252 T; 0 U; 0 Other;
QY	Query Match	51.2%; Score 1244.2; DB 2; Length 1263;
QY	Best Local Similarity	99.8%; Pred. No. 4e-274;
QY	Matches 1246; Conservative	0; Mismatches 3; Indels 0; Gaps 0;
DB	1181 GAGCAGCTCATCAACCCCTTTGGAGAGATGATGATGATTTTGGACCAACTGTGATGTC	1240
DB	1 GAGCAGCTCATCAACCCCTTTGGAGAGATGATGATGATTTTGGACCAACTGTGATGTC	60
QY	1241 GACGAGAAATTTGGACGAGTGTCCCTTTGGCTGTGGATAGATGACCAAGACCTGCTCCG	1300
DB	61 GACGAGAAATTTGGACGAGTGTCCCTTTGGCTGTGGATAGATGACCAAGACCTGCTCCG	120

CC chromosomes or to map related gene positions, to compare with endogenous
CC DNA sequences in patients to identify potential genetic disorders, as
CC probes to hybridise and discover novel related DNA sequences, as a source
CC of information to derive PCR primers for genetic fingerprinting, to raise
CC anti-protein antibodies and in gene therapy. The proteins can be used to
CC raise antibodies or to elicit another immune response, as reagents in
CC assays designed to quantitatively determine levels of the protein in
CC biological fluids, as markers for tissues in which the corresponding
CC protein is preferentially expressed and to treat autoimmune disorders
CC (e.g. multiple sclerosis, systemic lupus erythematosus, insulin dependent
CC diabetes mellitus or graft-versus-host disease), anaemia, periodontal
CC disorders, bone fractures, cartilage damage, central nervous system
CC disorders (e.g. Alzheimer's disease or Parkinson's disease) and cancers.
CC The proteins and polynucleotides are also useful as nutritional sources
CC or supplements, e.g. as carbon, nitrogen or carbohydrate sources. This
CC sequence represents a human polynucleotide of the invention.

XX Sequence 1238 BP, 387 A, 336 C, 266 G, 249 T, 0 U, 0 Other;

Query Match 41.5%; Score 1008.2; DB 9; Length 1238;
Best Local Similarity 97.9%; Pred. No. 3.5e-220;
Matches 1223; Conservative 0; Mismatches 3; Indels 23; Gaps 19;

QY 1181 GAGCAGCTCATCAACCCCTTTGAGAGAGATGATGATTTTGAAGCACTGATTTGTC 1240
DB 1 GAGCAGCTCATCAACCCCTTTGAGAGAGATGATGATTTTGAAGCACTGATTTGTC 60
QY 1241 GACAGGAATTTGACAGTGTCCCTGTGGCTGTGATGATGATTCACAGACTGCTCGG 1300
DB 61 GACAGGAATTTGACAGTGTCCCTGTGGCTGTGATGATGATTCACAGACTGCTCGG 119
QY 1301 ATGAGAGCCGACATGTATCTGGAATPAAAGCCGACAGCCGCCCTTACAGACTGCTCC 1360
DB 120 ATGAGAGCCGACATGTATCTGGAATPAAAGCCGACAGCCGCCCTTACAGACTGCTCC 178
QY 1361 GCCAGTTCGCTCGAGAGCTCTTTATGGGCTTCACATTCACATGAGCTTGAACAAAGAG 1420
DB 179 GCCAGTTCGCTCGAGAGCTCTTTATGGGCTTCACATTCACATGAGCTTGAACAAAGAG 237
QY 1421 GAGATGAGTTCACAGCCCAATCAGAGAGACGAGAGATGCTCAGCTGTGATCTTTGGC 1480
DB 238 GAGATGAGTTCACAGCCCAATCAGAGAGACGAGAGATGCTCAGCTGTGATCTTTGGG 296
QY 1481 CGCTTCCTCAAGGCTGAGTCCCATATGATCAACATCCCTCCAGAGGCAAACTCAAGAGCAAA 1540
DB 297 CGCTTCCTCAAGGCTGAGTCCCATATGATCAACATCCCTCCAGAGGCAAACTCAAGAGCA 355
QY 1541 CTACTGTGGCCCAAGAGGGAATCCCTTCTCCACAGAGGCTGCCCCAAAACACAGAGCA 1600
DB 356 CTACTGTGGCCCAAGAGGGAATCCCTTCTCCACAGAGGCTGCCCCAAAACACAGAGGC 414
QY 1601 GCCAAACAGAACTTTAGGGGCCAGAGACAAAGAGCTTGAAGCTTGAAGCTGTGGAC 1660
DB 415 GCCAAACAGAACTTTAGGGGCCAGAGACAAAGAGCTTGAAGCTTGAAGCTGTGGAG 473
QY 1661 GCCTTCAAGTGTGGCCCACTGATCAGAGGCGAGGCTCTCAAGGCCCCACAGAGCGCC 1720
DB 474 GCCTTCAAGTGTGGCCCACTGATCAGAGGCGAGGCTCTCAAGGCCCCACAGAGCG-CC 532
QY 1721 CTCAGCCCCACTCCCATGTTCTTCCCTTAGAACCATCAGCGCGTCAAGCTTCAACAGT 1780
DB 533 CTCAGCCCCACTCCCATGTTCTTCCCTTAGAACCATCAGAGCGCGTCAAGCTTCAACAG 591
QY 1781 GTCACAGGATAGACACCAAGACAAAGAGCTTAAAGCTGTGAGTTCTGGGGCCACAGAAA 1840
DB 592 GTCACAGGATAGACACCAAGACAAAGAGCTTAAAGCTGTGAGTTCTGGGGCCACAG-AA 650
QY 1841 AGTTTGAATGCTCTCAGAGAGGATGGGGCTTGAATGAGACCCGAGAGATCTCTGAA 1900
DB 651 AGTTTGAATGCTCTCAGAGAGGATGGGGCTTGAATGAGACCCGAGAGATCTCT-A 709
QY 1901 GTGAGAGAGAAACTGTGAGATTTAACTGACGATATGACAGAGATCCCGAAATGAC 1960

DB 710 GTGAGAGAGAAACTGTGAGATTTAACTGACGATATGCCAGAGATCCCGAAATGAC- 768
QY 1961 CTCAGAGAACTTTGGAAACATTCACACCAACCAATACACATCACTACATCAAGATCAGATG 2020
DB 769 CTCAGAGAACTTTGGAAACATTCACACCAACCAATACACATCACTACATCAAGATCAGAT 827
QY 2021 GATCCTTATTTGGGCTTTGGAAACAGGAGATGAGACATTCCTAACCTGCTTCTTAATGG 2080
DB 828 GATCCTTATTTGGGCTTTGGAAACAGGAGATGAGACATTCCTAACCTGCTTCTTAATG 886
QY 2081 GGAATGCTTGGCCAGCAGAGTCTTCACTGTGTGTAACACAGACAGACATGATTCAGTCA 2140
DB 887 GGAATGCTTGGCCAGCAGAGTCTTCACTGTGTGTAACACAGACAGACATGATTCAGTTC 945
QY 2141 CAGCATACAGCTGTGCCACATGAAGAACGAGTCTTCAACAGGCTGAATCAATGTTA 2200
DB 946 CAGCATACAGCTGTGCCACATGAAGAACGAGTCTTCAACAGGCTGAATCAATG--C 1003
QY 2201 GCTTAAATGATTAATAATCCAGACTACTTCACTTAAATGCTTTTAAATGATTAATAAACT 2260
DB 1004 GCTTAAATGATTAATAATCCAGACTACTTCACTTAAATGCTTTTAAATGATTAATAAA-- 1061
QY 2261 GTGAAAGCTAGACCTGAACCATTTGGAACATTTAACTCAGACTCTGAGATTCAGAGTCGGA 2320
DB 1062 GTGAAAGCTAGACCTGAACCATTTGGAACATTTAACTCAGACTCTGAGATTCAGAGTCG 1119
QY 2321 ACCCTTAGTCTATCTGAATCCAAGACGACACCTTAGTACTGCCAAACTAATGA 2380
DB 1120 ACCCTTAGTCTATCTGAATCCAAGACGACACCTTAGTACTGCCAAACTAAT-- 1177
QY 2381 GTTAAATTAATTAATAAATCTGTTAAAAA 2429
DB 1178 GTTAAATTAATTAATAAATCTGTTAAAAA 1226

RESULT 8
AAZ21229
ID AAZ21229 standard; cDNA; 1916 BP.
XX
AC AAZ21229;
XX
DT 22-NOV-1999 (first entry)
XX
DE Mouse CG1CE cDNA sequence.
XX
XX CG1CE; Best's macular dystrophy; mutation; diagnosis; detection; BMD;
XX age-related macular dystrophy; ss.
OS Mus sp.
XX
FH Key
FT CDS Location/Qualifiers
FT 11..166
FT /*tag= a
FT /product= "CG1CE protein"
XX
XX MO9943695-A1.
XX
XX PD 02-SEP-1999.
XX
XX PF 22-FEB-1999; 99WO-US003790.
XX
XX PR 25-FEB-1998; 98US-0075941P.
XX PR 18-DEC-1998; 98US-0112926P.
XX
XX PA (MERI-) MERCK & CO INC.
XX PA (UNIV-) UNIV UPSALA.
XX
XX PI Petrunkin K, Caskey CT, Metzker M, Madellius C;
XX WP1; 1999-540560/45.
XX DR P-PSDB; M4V29955.
XX
XX Human and mouse polynucleotides encoding CG1CE polypeptides.


```

XX Key Location/Qualifiers
FH Intron 1..1513
FT /*tag= a
FT /number= 1
FT 1514..1570
FT /*tag= b
FT /number= 1
FT 1571..1578
FT /*tag= c
FT /number= 2
FT 1579..1586
FT /*tag= d
FT /number= 2
FT 1587..1594
FT /*tag= e
FT /number= 2
FT 1595..1602
FT /*tag= f
FT /number= 3
FT 1603..1610
FT /*tag= g
FT /number= 3
FT 1611..1618
FT /*tag= h
FT /number= 4
FT 1619..1626
FT /*tag= i
FT /number= 4
FT 1627..1634
FT /*tag= j
FT /number= 5
FT 1635..1642
FT /*tag= k
FT /number= 5
FT 1643..1650
FT /*tag= l
FT /number= 6
FT 1651..1658
FT /*tag= m
FT /number= 6
FT 1659..1666
FT /*tag= n
FT /number= 7
FT 1667..1674
FT /*tag= o
FT /number= 7
FT 1675..1682
FT /*tag= p
FT /number= 8
FT 1683..1690
FT /*tag= q
FT /number= 8
FT 1691..1698
FT /*tag= r
FT /number= 9
FT 1699..1706
FT /*tag= s
FT /number= 9
FT 1707..1714
FT /*tag= t
FT /number= 10
FT 1715..1722
FT /*tag= u
FT /number= 10
FT 1723..1730
FT /*tag= v
FT /number= 11
FT 1731..1738
FT /*tag= w
FT /number= 11
XX

```

```

PN WO9943695-A1.
XX 02-SEP-1999.
XX 22-FEB-1999; 99WO-US003790.
XX 25-FEB-1998; 98US-0075941P.
XX 18-DEC-1998; 98US-0112926P.
XX (MERI ) MERCK & CO INC.
XX (UTUP-) UNIV UPPSALA.
XX Petrukhin K, Caskey CT, Metzker M, Wadelius C;
XX MPI; 1999-540560/45.
XX P-PSDB; AAY29953.
XX Human and mouse polynucleotides encoding CGICE polypeptides.
XX Claim 2; Fig 1; 67pp; English.
XX The present sequence represents the human CGICE gene, which when mutated
XX is responsible for Best's macular dystrophy (BMD). Polynucleotides
XX encoding CGICE are useful for diagnosing whether a patient carries a
XX mutation in the CGICE gene. Normal and mutated CGICE proteins are useful
XX for identifying activators and/or inhibitors of these proteins, in order
XX to treat BMD. The CGICE gene offers a simpler and cheaper method of
XX diagnosing BMD without the need for the presence of the patient. The gene
XX may also be useful to discovering the genetic cause of age-related
XX macular dystrophy
XX
XX Sequence 16125 BP; 3988 A; 4175 C; 4215 G; 3726 T; 0 U; 21 Other;
SQ
Query Match 26.4%; Score 642.2; DB 2; Length 16125;
Best Local Similarity 99.5%; Pred. No. 4e-136;
Matches 644; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1405 CAGCTGAACAAGAGAGATGAGATCCAGCCCAATCAGAGAGAGAGAGATGCTCA 1464
DB 13916 CAGCTGAACAAGAGAGATGAGATCCAGCCCAATCAGAGAGAGAGATGCTCA 13975
QY 1465 CGCTGGCATCTTTGGCGCTTCTAGGCTGCACTCCCATGATCAGCATCTCCAGGGC 1524
DB 13976 CGCTGGCATCTTTGGCGCTTCTAGGCTGCACTCCCATGATCAGCATCTCCAGGGC 14035
QY 1525 AAATCTCAAGACCAAACTACTGTGCGCCCAAGAGGGAATCCTTCTCCAGAGGCTGCC 1584
DB 14036 AAATCTCAAGACCAAACTACTGTGCGCCCAAGAGGGAATCCTTCTCCAGAGGCTGCC 14095
QY 1585 CAAAACACAGAGCAGCCAAAGCAAGACGTTAGGGGGCCAGAGACAAAGAGGCTGGAA 1644
DB 14096 CAAAACACAGAGCAGCCAAAGCAAGACGTTAGGGGGCCAGAGACAAAGAGGCTGGAA 14155
QY 1645 GCTTAAAGCTGTGAGCGCTTCAAGTCTGCGCCCACTGATCAGAGGCGCACTACTACG 1704
DB 14156 GCTTAAAGCTGTGAGCGCTTCAAGTCTGCGCCCACTGATCAGAGGCGCACTACTACG 14215
QY 1705 TGCCCCACAGAGCGCCCTCAGCCCACTCCATGTTCTTCCCTAGAACCATCAGCGCC 1764
DB 14216 TGCCCCACAGAGCGCCCTCAGCCCACTCCATGTTCTTCCCTAGAACCATCAGCGCC 14275
QY 1765 GTCAAAGCTTCAGAGTGCACAGGATAGACACCAAGAGCAAAAGCTTAAAGACTGTGAG 1824
DB 14276 GTCAAAGCTTCAGAGTGCACAGGATAGACACCAAGAGCAAAAGCTTAAAGACTGTGAG 14335
QY 1825 TTCTGGGGCCCAAGAAAGTTTGAATTGCTCTCAGAGAGCGATGGGCTTGAAGACA 1884
DB 14336 TTCTGGGGCCCAAGAAAGTTTGAATTGCTCTCAGAGAGCGATGGGCTTGAAGACA 14395
QY 1885 CCCAAGATATCTCAAGTGAAGAGAAAGCTGTGAGTTTAACTGACGGATATGCCAGA 1944
DB 14396 CCCAAGATATCTCAAGTGAAGAGAAAGCTGTGAGTTTAACTGACGGATATGCCAGA 14455

```

QY 1945 GATCCCGAAATACCTCAAGACCTTTGGAACATCCACCAACATACACTAC 2004
DB 14456 GATCCCGAAATACCTCAAGACCTTTGGAACATCCACCAACATACACTAC 14515
QY 2005 ACTCAAGATCATGATGATCTTATGGGCTTGGAAACAGGAG 2051
DB 14516 ACTCAAGATCATGATGATCTTATGGGCTTGGAAACAGGCTCG 14562

RESULT 10
ABAI4559
ID ABAI4559 standard; DNA; 16650 BP.
XX
AC ABAI4559;
XX
DT 23-JAN-2002 (first entry)
XX
DE Human nervous system related polynucleotide SEQ ID NO 6890.
XX
KW Human; nocotropic; neuroprotective; cytosolic; dermatological; virucide;
KW immunosuppressive; antihistaminic; anti-HIV; antibacterial; vulnery;
KW antiparkinsonian; antischistosomal; antianemic; antidiabetic; cancer;
KW antidiabetic; hepatocellular; cerebroprotective; antineoplastic;
KW antidiabetic; antidiabetic; antidiabetic; antidiabetic; antidiabetic;
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
XX
OS Homo sapiens.
XX
PN WO200159063-A2.
XX
PD 16-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001334.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 11-JUL-2000; 2000US-0216880P.
PR 14-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226861P.
PR 22-AUG-2000; 2000US-0226868P.
PR 23-AUG-2000; 2000US-0227182P.
PR 30-AUG-2000; 2000US-0227009P.
PR 01-SEP-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.

PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 20-OCT-2000; 2000US-0242221P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.

	17-NOV-2000;	2000US-0249214P	
PR	17-NOV-2000;	2000US-0249215P	
PR	17-NOV-2000;	2000US-0249216P	
PR	17-NOV-2000;	2000US-0249217P	
PR	17-NOV-2000;	2000US-0249218P	
PR	17-NOV-2000;	2000US-0249219P	
PR	17-NOV-2000;	2000US-0249220P	
PR	17-NOV-2000;	2000US-0249221P	
PR	17-NOV-2000;	2000US-0249222P	
PR	17-NOV-2000;	2000US-0249223P	
PR	17-NOV-2000;	2000US-0249224P	
PR	17-NOV-2000;	2000US-0249225P	
PR	17-NOV-2000;	2000US-0249226P	
PR	17-NOV-2000;	2000US-0249227P	
PR	17-NOV-2000;	2000US-0249228P	
PR	17-NOV-2000;	2000US-0249229P	
PR	17-NOV-2000;	2000US-0249230P	
PR	01-DEC-2000;	2000US-0250391P	
PR	01-DEC-2000;	2000US-0251160P	
PR	05-DEC-2000;	2000US-0251030P	
PR	05-DEC-2000;	2000US-0251988P	
PR	05-DEC-2000;	2000US-0256719P	
PR	06-DEC-2000;	2000US-0251479P	
PR	08-DEC-2000;	2000US-0251856P	
PR	08-DEC-2000;	2000US-0251868P	
PR	08-DEC-2000;	2000US-0251989P	
PR	08-DEC-2000;	2000US-0251989P	
PR	11-DEC-2000;	2000US-0251990P	
PR	05-JAN-2001;	2000US-0259678P	
PA	(HUMA-) HUMAN GENOME SCI INC.		
PI	Rosen CA, Barash SC, Ruben SM;		
XX	WPI; 2001-541565/60.		
DR	Nucleic acids encoding 3224 human nervous system antigen polypeptides,		
XX	useful for preventing, diagnosing and/or treating nervous system cancers		
PT	and metastases.		
PT			
PS	Disclosure; SEQ ID NO 6890; 1701bp + Sequence Listing; English.		
XX			
XX	The invention relates to novel genes (ABAI1004-ABA21534) and proteins		
CC	(ABAI1678-ABAI18001) useful for preventing, treating or ameliorating		
CC	medical conditions e.g. by protein or gene therapy. The genes are		
CC	isolated from a range of human tissues disclosed in the specification.		
CC	The nucleic acids, proteins, antibodies and (ant)agonists are useful in		
CC	the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and		
CC	ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,		
CC	breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune		
CC	disorders e.g. Addison's disease, allergies, autoimmune hemolytic		
CC	diseases e.g. Crohn's disease, diabetes mellitus, Crohn's disease,		
CC	anemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,		
CC	multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)		
CC	cardiovascular disorders such as myocardial ischemias; (d) wound healing		
CC	(e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)		
CC	infectious diseases such as viral, bacterial, fungal and parasitic		
CC	infections. Note: The sequence data for this patent did not form part of		
CC	the printed specification, but was obtained in electronic format directly		
CC	from WIPO at ftp.wipo.int/pub/published_pct_sequences		
XX			
SQ	Sequence 16650 BP; 4168 A; 4271 C; 4308 G; 3903 T; 0 U; 0 Other;		
	Query Match 26.4%; Score 642.2; DB 5; Length 16650;		
	Best Local Similarity 99.5%; Pred. No. 4.1e-136;		
	Matches 644; Conservative 0; Mismatches 3; Indels 0; Gaps 0;		
Dh			
	1405 CAGCCTGAACAAAGAGGATGGAGTTCACGCCCATCGAGGAGCAAGAGGATGTCTCA	1464	
	11895 CAGCCTGAACAAAGAGGATGGAGTTCACGCCCATCGAGGAGCAAGAGGATGTCTCA	11954	
Oy			
	1465 CGCTGGCATCTTGCGCGCTTCTCTAGGCCTGACATCCATGATCACATCTTCCAGGCG	1524	
Dh			
	11955 CGCTGGCATCTTGCGCGCTTCTCTAGGCCTGACATCCATGATCACATCTTCCAGGCG	12014	
Oy			
	1525 AAATCTCAAGACCAAATCTGTTGGCCCAAGAGGAAATCCCTTCTCAAGAGGCGTGGC	1584	
Dh			
	12015 AAATCTCAAGACCAAATCTGTTGGCCCAAGAGGAAATCCCTTCTCAAGAGGCGTGGC	12074	

Oy	1585	CAAAAACCAAAAGCGACCCAAAAGAAACGTTAGGGGCGAGAAAGCAAAAGGCGTGGAA	1644
Db	12075	CAAAAACCAAAAGCGACCCAAAAGAAACGTTAGGGGCGAGAAAGCAAAAGGCGTGGAA	12134
Oy	1645	GCTTAAGGCTGTGGACGCGCTTCAAGTCTGGGCCCATGTATCAGAGGCGCAGGCTACTACAG	1704
Db	12135	GCTTAAGGCTGTGGACGCGCTTCAAGTCTGGGCCCATGTATCAGAGGCGCAGGCTACTACAG	12194
Oy	1705	TGCCCCACAGAGCGCCCTCAGGCCCACTCCCATGTTCTTCCCTTAGAACCATCAGCGCC	1764
Db	12195	TGCCCCACAGAGCGCCCTCAGGCCCACTCCCATGTTCTTCCCTTAGAACCATCAGCGCC	12254
Oy	1765	GTCAAAGCTTCACAGTGTCAAGGCGATAGACACCAAGACAAAGAGCTTAAAGACTGTGAG	1824
Db	12255	GTCAAAGCTTCACAGTGTCAAGGCGATAGACACCAAGACAAAGAGCTTAAAGACTGTGAG	12314
Oy	1825	TTCTGGGGCCAAAGAAAAGTTTGAATTGCTCTCAGAGACGATGGGGCCTTGATGGAGCA	1884
Db	12315	TTCTGGGGCCAAAGAAAAGTTTGAATTGCTCTCAGAGACGATGGGGCCTTGATGGAGCA	12374
Oy	1885	CCCAAGAGTATCTCAAGTGAAGAGGAAAACTGTGAGTTTAACCTGACGATATGCGAGA	1944
Db	12375	CCCAAGAGTATCTCAAGTGAAGAGGAAAACTGTGAGTTTAACCTGACGATATGCGAGA	12434
Oy	1945	GATCCCGGAAATCACCCTCAAAGAACCTTTGGAACATACCAACCAATACACACTAC	2004
Db	12435	GATCCCGGAAATCACCCTCAAAGAACCTTTGGAACATACCAACCAATACACACTAC	12494
Oy	2005	ACTCAAGATCACATGGATCCTTATTGGGCTTTGAAAACAGGAGATG	2051
Db	12495	ACTCAAGATCACATGGATCCTTATTGGGCTTTGAAAACAGGATCG	12541
RESULT 11			
ID	ABAI4557	standard; DNA; 18530 BP.	
AC	ABAI4557;		
DT	23-JAN-2002	(first entry)	
DE	Human nervous system related polynucleotide SEQ ID NO 6888.		
XX			
KM	Human, noctropic; neuroprotective; cytoskeletal; dermatological; virucide;		
KM	Immunosuppressive; anti-inflammation; anti-HIV; antibacterial; vulnerary;		
KM	anti-parkinsonian; anti-asthma; anti-anemic; anti-arthritic; cancer;		
KM	anti-neumatic; hepatocytic; cerebroprotective; anti-inflammation;		
KM	anti-asthma; anti-diabetic; anticancer; anticonvulsant; antifungal;		
KM	anti-parasitic; cardiac; immune disorder; cardiovascular disorder;		
KM	neurological disease; infection; nephrotoxic; gene therapy; vaccine; ds.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200159063-A2.		
XX			
PD	16-AUG-2001.		
XX			
PF	17-JAN-2001; 2001WO-US001334.		
XX			
PR	31-JAN-2000; 2000US-0179065P.		
PR	04-FEB-2000; 2000US-0180628P.		
PR	24-FEB-2000; 2000US-0184664P.		
PR	02-MAR-2000; 2000US-0186350P.		
PR	16-MAR-2000; 2000US-0189674P.		
PR	17-MAR-2000; 2000US-0190076P.		
PR	18-APR-2000; 2000US-0198123P.		
PR	19-MAY-2000; 2000US-020515P.		
PR	07-JUN-2000; 2000US-0209467P.		
PR	28-JUN-2000; 2000US-0214886P.		
PR	30-JUN-2000; 2000US-0215135P.		
PR	07-JUL-2000; 2000US-0216647P.		
PR	07-JUL-2000; 2000US-0216880P.		

RESULT 11
ABAI4557
ID ABAI4557 standard; DNA; 18530 BP.
XX
XX ABAI4557;
XX
DT 23-JAN-2002 (first entry)
XX
DE Human nervous system related polynucleotide SEQ ID NO 6888.
XX
KW Human; nootropic; neuroprotective; cytosolic; dermatological; virucide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;
KW antitumoric; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antileucy; anticonvulsant; antifungal;
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
KW neurological disease; infection; neurotropic; gene therapy; vaccine; ds.
XX
OS Homo sapiens.
XX
XX
PN WO200159063-A2.
XX
PD 16-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001334.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.

PR	21-SEP-2000;	2000US-0234274P
PR	25-SEP-2000;	2000US-0234597P
PR	26-SEP-2000;	2000US-0234598P
PR	26-SEP-2000;	2000US-0235484P
PR	27-SEP-2000;	2000US-02355834P
PR	27-SEP-2000;	2000US-0235536P
PR	29-SEP-2000;	2000US-0236327P
PR	29-SEP-2000;	2000US-0236367P
PR	29-SEP-2000;	2000US-0236368P
PR	29-SEP-2000;	2000US-0236369P
PR	29-SEP-2000;	2000US-0236370P
PR	02-OCT-2000;	2000US-0236802P
PR	02-OCT-2000;	2000US-0237037P
PR	02-OCT-2000;	2000US-0237038P
PR	02-OCT-2000;	2000US-0237039P
PR	02-OCT-2000;	2000US-0237039P
PR	13-OCT-2000;	2000US-0237040P
PR	13-OCT-2000;	2000US-0239353P
PR	13-OCT-2000;	2000US-0239353P
PR	20-OCT-2000;	2000US-0240960P
PR	20-OCT-2000;	2000US-0241785P
PR	20-OCT-2000;	2000US-0241786P
PR	20-OCT-2000;	2000US-0241787P
PR	20-OCT-2000;	2000US-0241808P
PR	20-OCT-2000;	2000US-0241808P
PR	20-OCT-2000;	2000US-0241826P
PR	01-NOV-2000;	2000US-0242221P
PR	01-NOV-2000;	2000US-0244617P
PR	08-NOV-2000;	2000US-0246474P
PR	08-NOV-2000;	2000US-0246475P
PR	08-NOV-2000;	2000US-0246476P
PR	08-NOV-2000;	2000US-0246477P
PR	08-NOV-2000;	2000US-0246477P
PR	08-NOV-2000;	2000US-0246528P
PR	08-NOV-2000;	2000US-0246532P
PR	08-NOV-2000;	2000US-0246532P
PR	08-NOV-2000;	2000US-0246610P
PR	08-NOV-2000;	2000US-0246611P
PR	08-NOV-2000;	2000US-0246613P
PR	17-NOV-2000;	2000US-0249207P
PR	17-NOV-2000;	2000US-0249208P
PR	17-NOV-2000;	2000US-0249209P
PR	17-NOV-2000;	2000US-0249210P
PR	17-NOV-2000;	2000US-0249211P
PR	17-NOV-2000;	2000US-0249212P
PR	17-NOV-2000;	2000US-0249213P
PR	17-NOV-2000;	2000US-0249214P
PR	17-NOV-2000;	2000US-0249215P
PR	17-NOV-2000;	2000US-0249216P
PR	17-NOV-2000;	2000US-0249217P
PR	17-NOV-2000;	2000US-0249218P
PR	17-NOV-2000;	2000US-0249245P
PR	17-NOV-2000;	2000US-0249245P
PR	17-NOV-2000;	2000US-0249264P
PR	17-NOV-2000;	2000US-0249265P
PR	17-NOV-2000;	2000US-0249297P
PR	17-NOV-2000;	2000US-0249299P
PR	17-NOV-2000;	2000US-0249300P
PR	01-DEC-2000;	2000US-0250391P
PR	01-DEC-2000;	2000US-0250391P
PR	05-DEC-2000;	2000US-0251030P
PR	05-DEC-2000;	2000US-0251030P
PR	05-DEC-2000;	2000US-0251671P
PR	06-DEC-2000;	2000US-0251671P
PR	08-DEC-2000;	2000US-0251856P
PR	08-DEC-2000;	2000US-0251856P
PR	08-DEC-2000;	2000US-0251863P
PR	08-DEC-2000;	2000US-0251989P
PR	08-DEC-2000;	2000US-0251990P

PR	11-DEC-2000; 2000US-0254097P.
PR	05-JAN-2001; 2001US-0259678P.
XX	(HUMA-) HUMAN GENOME SCI INC.
PA	
PI	Rosen CA, Barash SC, Ruben SM,
PT	WPI; 2001-541565/60.
DR	
XX	
XX	
PS	Nucleic acids encoding 3224 human nervous system antigen polypeptides, PT useful for preventing, diagnosing and/or treating nervous system cancers PR and metastases. disclosures.
SS	
XX	Sequence; SEQ ID NO 6889; 1701bp + Sequence Listing; English.
CC	The invention relates to novel genes (ABA11004-ABA21534) and proteins
CC	(ABA14678-ABA18001) useful for preventing, treating or ameliorating
CC	medical conditions e.g. by protein or gene therapy. The genes are
CC	isolated from a range of human tissues disclosed in the specification.
CC	The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC	the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC	ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC	breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC	disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC	anemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC	multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC	cardiovascular disorders such as myocardial ischaemias; (d) wound healing
CC	; (e) neurological diseases such as viral, bacterial, fungal and parasitic
CC	infectious diseases such as viral, bacterial, fungal and parasitic
CC	infections. Note: The sequence data for this patent did not form part of
CC	the printed specification, but was obtained in electronic format directly
CC	from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX	
XX	
SQ	Sequence 18537 BP; 4519 A; 4800 C; 4983 G; 4235 T; 0 U; 0 Other;
Query Match	26.4%; Score 640.6; DB 5; Length 18537;
Best Local Similarity	99.4%; Pred. No. 9.8e-136;
Matches 643; Conservative	0; Mismatches 4; Indels 0; Gaps 0;
OY	1405 CAGCCTGAACAAAGGAGATGGAGTTCACGCCCAATCACAGAGAAGAGATCTCTCA
Db	11901 CAGCCTGAACAAAGGAGATGGAGTTCACGCCCAATCACAGAGAAGAGATCTCTCA
OY	1465 CGCTGCATCATTTGGCGGCTTCTTAGGCTTGCAATCATCATCTCTCCAGGGC
Db	11961 CGCTGCATCATTTGGCGGCTTCTTAGGCTTGCAATCATCATCTCTCCAGGGC
OY	1525 AAACTCAGAGCCCAAATCATCTGTGGCCCCAAGAGAGAAATCCCTTCTCCAAGAGGCTGGC
Db	12021 AAACTCAGAGCCCAAATCATCTGTGGCCCCAAGAGAGAAATCCCTTCTCCAAGAGGCTGGC
OY	1585 CAAAAACCAAGGACGCCCAAGAACGTTAGGGCCAGAGAAACAACAAAGGCTGGAA
Db	12081 CAAAAACCAAGGACGCCCAAGAACGTTAGGGCCAGAGAAACAACAAAGGCTGGAA
OY	1645 GCTTAAGGCTGTGGAGCGCTTCAAGTCTGGCCCATGTATCAAGAGCCAGCTACTAAG
Db	12141 GCTTAAGGCTGTGGAGCGCTTCAAGTCTGGCCCATGTATCAAGAGCCAGCTACTAAG
OY	1705 TGCCCAACAGACGCCCTCAGGCCCACTCCAAAGTTCTTCCCCTTAAGAACCATACAGGCC
Db	12201 TGCCCAACAGACGCCCTCAGGCCCACTCCAAAGTTCTTCCCCTTAAGAACCATACAGGCC
OY	1765 GTCAAAAGCTTCAACAGTGTCAACAGCATAGACCAAAAGCAAAGCTTAAAGACTGTGAG
Db	12261 GTCAAAAGCTTCAACAGTGTCAACAGCATAGACCAAAAGCAAAGCTTAAAGACTGTGAG
OY	1825 TTCTGGGGCCCAAGAAAAGTTTTGAATTGGTCTCAAGAGCGAATGGGCTTGAATGAGACA
Db	12321 TTCTGGGGCCCAAGAAAAGTTTTGAATTGGTCTCAAGAGCGAATGGGCTTGAATGAGACA
OY	1885 CCAGAAGTATCTCAAGTGAAGAGAAAACTGTGAGTTTAACTGACGATATGACAGA
Db	1944

Db	12381	CCCAAGATCTCAAGTGAAGAGAGAGAAAACGTGAGATTAACTGACGATATGCGAGA	1244
Qy	1945	GATCCCGGAAATACCTCTCAAGAAACCTTTGGAAACATCAACCAACTATACACTAC	2004
Db	12441	GATCCCGGAAATACCTCTCAAGAAACCTTTGGAAACATCAACCAACTATACACTAC	12500
Qy	2005	ACTCAAAAGATACATGATCCTCTTTATTTGGGCTTGGAAACAGGATG	2051
Db	12501	ACTCAAAAGATACATGATCCTCTTTATTTGGGCTTGGAAACAGGATCTG	12547
RESULT_13			
ID	ABL89697	standard; CDNA; 1717 BP.	
AC	ABL89697;		
DT	24-MAY-2002	(first entry)	
XX		Human polynucleotide SEQ ID NO 259.	
DE			
XX		Cytostatic; immunosuppressive; neutrotropic; neuroprotective; antiviral;	
KW		antiallergic; hepatotropic; antidiabetic; antiinflammatory; anticancer;	
KW		wulnery; anticonvulsant; antibacterial; antifungal; antiparasitic;	
KM		cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;	
XX		neurological disease; infection; human; secreted protein; gene; ss.	
XX			
OS		Homo sapiens.	
XX			
PN		WO200190304-A2.	
PD		29-NOV-2001.	
XX			
PP		18-MAY-2001; 2001WO-US016450.	
XX			
PR		19-MAY-2000; 2000US-0205515P.	
XX			
PA		(HUMA-) HUMAN GENOME SCI INC.	
PI		Blirge CE, Rosen CA;	
DR		WPI; 2002-122018/16.	
DR		P-PSDB; ABB85288.	
PT		Novel 1405 isolated polypeptides, useful for diagnosis, treatment and	
PT		prevention of neural, immune system, muscular, reproductive,	
PT		gastrointestinal, pulmonary, cardiovascular, renal and proliferative	
PT		disorders.	
XX			
PS		Claim 4; SEQ ID NO 259; 2081pp + Sequence Listing; English.	
XX			
CC		The invention relates to novel genes (ABL89449-ABL90853) and proteins	
CC		(ABB8940-ABB9044) useful for preventing, treating or ameliorating	
CC		medical conditions e.g., by protein or gene therapy. The genes are	
CC		isolated from a range of human tissues disclosed in the specification.	
CC		The nucleic acids, proteins, antibodies and (ant)agonists are useful in	
CC		the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and	
CC		ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,	
CC		breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune	
CC		disorders e.g. Addison's disease, allergies, autoimmune hemolytic	
CC		anemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,	
CC		multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)	
CC		cardiovascular disorders such as myocardial ischaemias; (d) wound healing	
CC		; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)	
CC		infectious diseases such as viral, bacterial, fungal and parasitic	
CC		infections. Note: The sequence data for this patent did not form part of	
CC		the printed specification, but was obtained in electronic format directly	
CC		from WIPO at ftp.wipo.int/pub/published_pct_sequences	
XX			
XX			
SEQ		Sequence 1717 BP; 464 A; 460 C; 417 G; 367 T; 0 U; 9 Other;	
Query Match	24.2%	Score 587.8; DB 6; Length 1717;	
Best Local Similarity	98.8%;	Pred. NO. 4.9e-124;	

	Matches	600; Conservative	2; Mismatches	4; Indels	1; Gaps	1;
QY	1445	GAAGACGAGGAGAGATGCTCAGGCTGGACATCATTTGGCCGCTCTTAGGCTCGAGTCCAT	1504			
Db	1	GAGGACGAGGAGGAGATGCTCAGGCTGGACATCATTTGGCCGCTCTTAGGCTCGAGTCCAT	60			
QY	1505	GATCACCATCTCTCCCAAGGGCAAACTCAAGGACCAAACTACTGTGGCCCAAGAGGAAATCC	1564			
Db	61	GATCACCATCTCTCCCAAGGGCAAACTCAAGGACCAAACTACTGTGGCCCAAGAGGAAATCC	120			
QY	1565	CTTCTCCACGAGGGGCTCCCAAAAAACCAAGGACGAGCCAAACAGATTAGGGGCGAG	1624			
Db	121	CTTCTCCACGAGGGGCTG-CCAAAMAACCAAGGACGCAAAACGATAGGGGCGAG	179			
QY	1625	GAAGACAAAGGCGCTGGAGCTTAGGCTGTGAGCGCTTCAAGTCTGGGCCCACTGTAT	1684			
Db	180	GAAGACAAAGGCGCTGGAGCTTAGGCTGTGAGCGCTTCAAGTCTGGGCCCACTGTAT	239			
QY	1685	CAGAGGCCAGGCTACTAGTAGGCCCAAGAGCGCCCTCAGCCCACTCCCATGTTCTTC	1744			
Db	240	CAGAGGCCAGGCTACTAGTAGGCCCAAGAGCGCCCTCAGCCCACTCCCATGTTCTTC	299			
QY	1745	CCCCTAGAAGCATCAGCGCCGCTCAAAAGCTTCAAGTGTCAAGGATAGACCAAAAGAC	1804			
Db	300	CCCCTAGAAGCATCAGCGCCGCTCAAAAGCTTCAAGTGTCAAGGATAGACCAAAAGAC	359			
QY	1805	AAAAAGCTTAAAGACTGTAGTTCGTGGGGCCAAAGAAAGTTTGAATTGCTCTCAGAGAGC	1864			
Db	360	AAAAAGCTTAAAGACTGTAGTTCGTGGGGCCAAAGAAAGTTTGAATTGCTCTCAGAGAGC	419			
QY	1865	GATGGGGGCTTGTATGTAGACACCCCAAGATATCTCAAGTAGAGAGAAACGTGTGAAGTTT	1924			
Db	420	GATGGGGGCTTGTATGTAGAGACACCCAGAAATATCTCAAGTAGAGAGAAACCGTGTGAAGTTT	479			
QY	1925	AACCTGACGGATGTATGCGCAGAGATCCGCCAAATCACTCAAGAAACCTTTGGAAACATCA	1984			
Db	480	AACCTGACGGATGTATGCGCAGAGATCCGCCAAATCACTCAAGAAACCTTTGGAAACATCA	539			
QY	1985	CCAAACCAACATACACACTACTCACTCAAGAGATCAATGATCCTTATTGGGCTTGGAAAA	2044			
Db	540	CCAAACCAACATACACACTACTCACTCAAGAGATCAATGATCCTTATTGGGCTTGGAAAA	599			
QY	2045	AGGGATG 2051				
Db	600	AGGTCTG 606				
RESULT 14						
ABAI4556						
ID	ABAI4556	standard; DNA; 7108 BP.				
XX	ABAI4556;					
AC	23-JAN-2002	(first entry)				
XX						
DE	Human nervous system related polynucleotide SEQ ID NO 6887.					
XX						
KM	Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;					
KM	immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;					
KM	antiparkinsonian; antispasmodic; antianemic; antirheumatic; cancer;					
KM	antihelminthic; hepatotropic; cerebroprotective; antiinflammatory;					
KM	antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;					
KM	antiparasitic; cardiac; immune disorder; cardiovascular disorder;					
KM	neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.					
XX						
OS	Homo sapiens.					
XX						
PN	WO200159063-A2.					
XX						
PD	16-AUG-2001.					
XX						
PF	17-JAN-2001; 2001WO-US001334.					
XX						

QY 403 TGCCTGAGCCCGACCGCCTCATGAGCCTGTGTGCGGCTTGTGCAAGCAAGACGAC 462
 Db 490 TGCCTGCTGCCCGACCGGCTCATGTGCGTGTGCGGCAACGTCACGGAACGACGAC 549
 QY 463 AAGGCGGCTGTGCTGCGGCGACAGCTCATCCGCTACGCCAACCTGGGCAACGCTCATCC 522
 Db 550 GCGGCGGCTTACCGGCGCACACTCATGCGCTACGAGGGCTCTCGGCGGCTGTCTCATCC 609
 QY 523 TCGCGAGGCTCAGACCGCAGCTCATCAGAGCCTTCCCGAGGCCAGCCAGCACTGTGTCAAG 582
 Db 610 TCGGCTCCGTCAGACACCGGCTTCAAGCGCTTCCCAACATAGACACGCTGTGAGG 669
 QY 583 CAGGCTTATGACTCCGCGAAGACCAAGCAGTTGAGAACTGAGCCTACACACACA 642
 Db 670 CTGGGTTATGACCGCGAGGAGGCAAGAGTTTGAAGAACTGACTCATCTACACACA 729
 QY 643 TGTCTGGGTCCTCGGCTGTGTGCTTGTCCAACTGTCAATGAAGCGTGTGAGGTC 702
 Db 730 AGTACTGGGTGCTCTGCGTGTGTTCACACTGGGCGCACAGGCCGACGAGGCGC 789
 QY 703 GAATCCGGGACCTATCTCTGCTCCAGAGCCTGTGACAGATGAAACACTTGTCTACTC 762
 Db 790 GCATCCGCGACAAACAGCGCCCTTAAGCTGCTCGAGAGCTGAATGTTTTCGCGGCA 849
 QY 763 AGTGTGACACCTGTATGCTTACGACTGATGATTAATCCACTGTGTATACACAGGTG 822
 Db 850 AATGTGAATGCTCTTCACTATGACTGATTAAGCTACCCCTGTGTACACGAGGTG 909
 QY 823 TGACTGTGGCGGTATACAGCTTCTCTGACTGTCTAGTTGGGCGGCAAGTTCTGAAC 882
 Db 910 TGACCATCGACATGTACAGCTACTTCTGTGCTTCAATGTGTGCGCAAGTCTGAGC 969
 QY 883 CAGCAAGGCTTACCTCGGCAATGAGCTGAACCTCGTGTGCGGCTTCAAGTCTGTC 942
 Db 970 CGGCTCAGGCTTCAAAAGACACAGACTGAGCTGTGTGCGCAATCTTCAACCTCTTGC 1029
 QY 943 AGTTCTTCTTCTATGTTGGCTGAGAGGTG 975
 Db 1030 AGTTCTTCTTCTAAGCGGCTGTGCTCAAGGTAG 1062

Search completed: March 25, 2005, 23:51:48
 Job time : 1275.21 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using BW model

Run on: March 26, 2005, 07:39:17 / Search time 1353.06 Seconds
(without alignments)
10698.156 Million cell updates/sec

Title: US-09-622-964A-4
Perfect score: 2429
Sequence: 1 caggagagccaccacagccta.....aaaaaaaaaaaaaaaaaaaaa 2429

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5552208 seqs, 297965951 residues

Total number of hits satisfying chosen parameters: 11104416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US10F_NEW_PUB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1541.8	63.5	1758	US-10-492-032-1	Sequence 1, Appl1
2	1244.2	51.2	1263	US-09-746-783-3	Sequence 3, Appl1
3	587.8	24.2	1717	US-10-264-237-259	Sequence 259, App
4	431.4	17.8	2137	US-10-492-032-3	Sequence 3, Appl1
5	431	17.7	1530	US-10-492-032-30	Sequence 30, Appl1
6	350.4	14.4	1326	US-10-071-766-10	Sequence 7, Appl1
7	339.4	14.0	1422	US-10-492-032-7	Sequence 1299, Ap
8	338.2	13.9	2500	US-10-108-260A-1299	Sequence 5, Appl1
9	336.8	13.9	2028	US-10-492-032-5	Sequence 19504, A
10	213	8.8	620	US-09-814-353-19504	Sequence 11070, A
11	189.2	7.8	1292	US-10-198-846-11070	

C 12	175	7.2	1198	9	US-09-880-107-2174	Sequence 2174, Ap
C 13	161.4	6.6	578	13	US-10-027-632-134530	Sequence 134530,
C 14	161.4	6.6	578	13	US-10-027-632-134530	Sequence 134530,
C 15	149.4	6.2	748	13	US-10-027-632-134529	Sequence 134529,
C 16	149.4	6.2	748	13	US-10-027-632-134529	Sequence 134529,
C 17	135.2	5.6	539	16	US-10-029-386-7583	Sequence 7583, Ap
C 18	134.2	5.5	235	16	US-10-029-386-7583	Sequence 21283, A
C 19	126.8	5.2	1350	9	US-09-768-826-16	Sequence 16, Appl
C 20	126.8	5.2	1350	18	US-10-874-484-16	Sequence 7441, Ap
C 21	124.4	5.1	2914	18	US-10-723-860-7441	Sequence 8303, Ap
C 22	124.4	5.1	2914	18	US-10-723-860-8303	Sequence 11346, A
C 23	116	4.8	853	14	US-10-198-846-11346	Sequence 7576, Ap
C 24	115.4	4.8	1954	18	US-10-723-860-7576	Sequence 2561, Ap
C 25	104	4.3	462	14	US-10-198-846-2561	Sequence 174961,
C 26	103.6	4.3	3186778	13	US-10-027-632-174961	Sequence 174961,
C 27	103.6	4.3	3186778	17	US-10-027-632-174961	Sequence 7843, Ap
C 28	102	4.2	615	14	US-10-198-846-7843	Sequence 195, Ap
C 29	98	4.0	507	10	US-09-764-872-195	Sequence 8, Appl1
C 30	91	3.7	106707	18	US-10-694-685-8	Sequence 243509,
C 31	90	3.7	624	13	US-10-027-632-243509	Sequence 243509,
C 32	90	3.7	624	17	US-10-027-632-243509	Sequence 427, App
C 33	90	3.7	267156	11	US-09-868-007A-427	Sequence 1, Appl1
C 34	90	3.7	684973	9	US-09-263-959-1	Sequence 5719, Ap
C 35	89.4	3.7	398287	18	US-10-741-601-5719	Sequence 17839, A
C 36	89.4	3.7	398287	19	US-10-741-601-5719	Sequence 599, App
C 37	89	3.7	31994	9	US-09-764-904-71	Sequence 71, Appl
C 38	89	3.7	31994	14	US-10-091-548-71	Sequence 599, App
C 39	89	3.7	31994	14	US-10-074-095-599	Sequence 599, App
C 40	89	3.7	31994	17	US-10-312-872-599	Sequence 11024, A
C 41	89	3.7	723	13	US-10-027-632-11024	Sequence 11024, A
C 42	88.4	3.6	723	17	US-10-027-632-11024	Sequence 148248,
C 43	88.4	3.6	727	13	US-10-027-632-148248	Sequence 148248,
C 44	88.4	3.6	727	13	US-10-027-632-148248	Sequence 148249,
C 45	88.4	3.6	727	13	US-10-027-632-148249	

ALIGNMENTS

RESULT 1
US-10-492-032-1
Sequence 1, Application US/10492032
Publication No. US20050049212A1
GENERAL INFORMATION:
APPLICANT: Steuerrnagel, Arnd
APPLICANT: Brommer, Gunter
APPLICANT: Fritsch, Rudiger
APPLICANT: Eulenberg, Karsten
TITLE OR INVENTION: Bestrophin and Bestrophin Homologous Proteins Involved in the
Regulation of Energy Homeostasis
FILE REFERENCE: 2923-606
CURRENT APPLICATION NUMBER: US/10/492,032
CURRENT FILING DATE: 2004-04-08
PRIOR APPLICATION NUMBER: PCT/EP02/11321
PRIOR FILING DATE: 2002-10-09
PRIOR APPLICATION NUMBER: EP01124059.5
NUMBER OF SEQ ID NOS: 37
SOFTWARE: PatentIn version 3.2
SEQ ID NO 1
LENGTH: 1758
TYPE: DNA
ORGANISM: Homo sapiens
US-10-492-032-1
Query Match 63.5%; Score 1541.8; DB 19; Length 1758;
Best Local Similarity 89.5%; Pred. No. 0;
Matches 1756; Conservative 0; Mismatches 2; Indels 203; Gaps 1;
Qy 105 ATGACATCACTTACACAGCAAGTGGCTAATGCCGCTTGGCTCTTCCGCTG 164
Db 1 ATGACATCACTTACACAGCAAGTGGCTAATGCCGCTTGGCTCTTCCGCTG 60

QY 165 CTGCTGCTGAGCGGAGAGATCTACAGCTGCTATATGCGAGTCTTAACTCTG 224
DB 61 CTGCTGCTGAGCGGAGAGATCTACAGCTGCTATATGCGAGTCTTAACTCTG 120
QY 225 CTGCTGCTGATCATCATCCCTTTATTTATAGCTGCGCCCTCAACGAAACAAACAGCTG 284
DB 121 CTGCTGCTGATCATCATCCCTTTATTTATAGCTGCGCCCTCAACGAAACAAACAGCTG 180
QY 285 ATGTTTGAAGAACTGATCTGTTATGCGAGCTTACATCATCTAGCTATCCCACTTCTTC 344
DB 181 ATGTTTGAAGAACTGATCTGTTATGCGAGCTTACATCATCTAGCTATCCCACTTCTTC 240
QY 345 GTGCTGAGCTTCTATCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 404
DB 241 GTGCTGAGCTTCTATCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 300
QY 405 CCGTGAGCCGAGCCGCTCATAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 464
DB 301 CCGTGAGCCGAGCCGCTCATAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
QY 465 GCGCGGCTGCTGCGGCGGAGCTCATCCGCTACGCGCAACCTGGGCAACGCTGCTATCTG 524
DB 361 GCGCGGCTGCTGCGGCGGAGCTCATCCGCTACGCGCAACCTGGGCAACGCTGCTATCTG 420
QY 525 CCGAGCTGAGCAGCGGAGCTCATAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 584
DB 421 CCGAGCTGAGCAGCGGAGCTCATAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
QY 585 GCGTTTATGATCTCGGAGCAACAGCAGTTGAGAGAACTGAGCTTACCAACCAACGATG 644
DB 481 GCGTTTATGATCTCGGAGCAACAGCAGTTGAGAGAACTGAGCTTACCAACCAACGATG 540
QY 645 TTCTGAGGCTGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 704
DB 541 TTCTGAGGCTGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
QY 705 ATCCGAGACCTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 764
DB 601 ATCCGAGACCTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
QY 765 TGTGAGACCTGATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 824
DB 661 TGTGAGACCTGATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
QY 825 ACTGTGCGGCTGATGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 884
DB 721 ACTGTGCGGCTGATGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
QY 885 GCCAAGGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 944
DB 781 GCCAAGGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
QY 945 TTCTTCTTCTGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1004
DB 841 TTCTTCTTCTGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 866
QY 1005 CATGSCAGAGGGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1064
DB 867 ----- 866
QY 1065 AAGGCTTCACGGGTAGAAAGCAGCCGAGCGTGGGCGCACACTGTAACTCCAGCTACT 1124
DB 867 ----- 866
QY 1125 CGGAGGCTGAGGAGAGATGCTTGAACCGGAGGCGGAGGTTGTTGAGGAGC 1184
DB 867 ----- 866
QY 1185 AGCTGATCAACCCCTTTGAGAGATGATGATTTTGAACAACCTGATTTGCGACA 1244
DB 878 AGCTGATCAACCCCTTTGAGAGATGATGATTTTGAACAACCTGATTTGCGACA 937

QY 1245 GGAATTTGAGAGTGTCCCTGTTGCTGCTGATGAGATGCAACAGAACCTGCTCGATGG 1304
DB 938 GGAATTTGAGAGTGTCCCTGTTGCTGCTGATGAGATGCAACAGAACCTGCTCGATGG 997
QY 1305 AGCCGAGCATATGATGAGAAATTAAGCCCGAGCAAGCCCCCTTACAGAGCTGCTTCCGCC 1364
DB 998 AGCCGAGCATATGATGAGAAATTAAGCCCGAGCAAGCCCCCTTACAGAGCTGCTTCCGCC 1057
QY 1365 AGTTCCGTGAGCCCTCTTTATGAGCTTCACCTTCAACATTCAGCTTAAACAAAGAGAGA 1424
DB 1058 AGTTCCGTGAGCCCTCTTTATGAGCTTCACCTTCAACATTCAGCTTAAACAAAGAGAGA 1117
QY 1425 TGAAGTTCAGCCCAATTAAGAGAGAGAGAGAGATGCTCAAGCTGATCATTTGAGCGCT 1484
DB 1118 TGAAGTTCAGCCCAATTAAGAGAGAGAGAGATGCTCAAGCTGATCATTTGAGCGCT 1177
QY 1485 TCTTAGGCTGAGCTTCCATGATCAACATCTCCAGGCGCAACCTGAAGAGCAACATAC 1544
DB 1178 TCTTAGGCTGAGCTTCCATGATCAACATCTCCAGGCGCAACCTGAAGAGCAACATAC 1237
QY 1545 TGTGAGCCAGAGAGAAATCCCTTCTCAAGAGGCTGCTGCAAAACCAAGAGAGCA 1604
DB 1238 TGTGAGCCAGAGAGAAATCCCTTCTCAAGAGGCTGCTGCAAAACCAAGAGAGCA 1297
QY 1605 AACAGAACCTTGAAGGAGCAGAGAGCAACAGAGCTGAGAGTTAAAGCTTGTGAGCGCT 1664
DB 1298 AACAGAACCTTGAAGGAGCAGAGAGCAACAGAGCTGAGAGTTAAAGCTTGTGAGCGCT 1357
QY 1665 TCAAGTCTGAGCCCACTGATTCAGAGAGCAGGCTACTAGAGTGGCCCAAGAGCTTCA 1724
DB 1358 TCAAGTCTGAGCCCACTGATTCAGAGAGCAGGCTACTAGAGTGGCCCAAGAGCTTCA 1417
QY 1725 GCGCCATCCCAATGTTCTTCCCTTGAAGACCATAGAGGCTTCAAGCTTCAAGTCTCA 1784
DB 1418 GCGCCATCCCAATGTTCTTCCCTTGAAGACCATAGAGGCTTCAAGCTTCAAGTCTCA 1477
QY 1785 CAGGCAATGAGCCCAACAAACAAAGCTTAAAGACTGATGATTTCTGAGGCGCAAGAAAGTT 1844
DB 1478 CAGGCAATGAGCCCAACAAACAAAGCTTAAAGACTGATGATTTCTGAGGCGCAAGAAAGTT 1537
QY 1845 TTGAATTTCTCTCAGAGAGCAGATGAGGCTTGAATGAGGCAACCAAGATTTCTCAAGTGA 1904
DB 1538 TTGAATTTCTCTCAGAGAGCAGATGAGGCTTGAATGAGGCAACCAAGATTTCTCAAGTGA 1597
QY 1905 GGAGGAGAACTGAGGATTTTAACTGACGAGATATGCCAGAGATCCCGGAAATCACTTCA 1964
DB 1598 GGAGGAGAACTGAGGATTTTAACTGACGAGATATGCCAGAGATCCCGGAAATCACTTCA 1657
QY 1965 AAGAACCCTTGAACAATCAACCAACCAACATACATCACTCAACCAAGATCAATGATGATC 2024
DB 1658 AAGAACCCTTGAACAATCAACCAACCAACATACATCACTCAACCAAGATCAATGATGATC 1717
QY 2025 CTTATTTGAGGCTTGAAGAACAGGATGAGACATTTCTTAA 2065
DB 1718 CTTATTTGAGGCTTGAAGAACAGGATGAGACATTTCTTAA 1758

RESULT 2

US-09-746-783-3
Sequence 3, Application US/09746783
Publication No. US20030044935A1
GENERAL INFORMATION:

APPLICANT: Jacobs, Kenneth

McCoY, John M.
Lavallee, Edward R.
Racie, Lisa A.

Treacy, Maurice
Spaulding, Vikki
Agostino, Michael J.
Howes, Steven H.

Fechele, Kim
TITLE OF INVENTION: SECRETED PROTEINS AND POLYPEPTIDES

ENCODING THEM

RESULT 4
US-10-492-032-3

QY 103 CGATGACATCATCTTACACAAAGCAAGAGGAGCTTAATGCCCGCTTAAAGCTCTTCTCCGCGC 162
 Db 210 CGATGACCGGTACCTTACACAAAGCCGAGTGGCCAAAGCCCGCTCTCGTGGCTTCTCCAGC 269
 QY 163 TGGTCTGTGCTGGCGGGCAGCATTTACAACCTGTGTATATGGCGAGTTCTTAATCTTCC 222
 Db 270 TGGCTGCTACTGTGGCGTGGAGAGATCTCAAACTCTGTGGGAGAGCTGCTCTGCTTCC 329
 QY 223 TGGCTGCTACTAATCATATCCGCTTATTTATAGGCTGGCCCTCAAGAGAAACAAGC 282
 Db 330 TTGGGTTTCAATAGGCGCTGAGTGTGCTTACCGTTTGTGCTGACCAAGGAGAGAAGC 389
 QY 283 TGAATGTTGAGAACTGACTCTGTATTTGCGACAGCTAATCATCACTCATCCCAATTTTCT 342
 Db 390 GCTACTTCAGAAAGCTTGTGATTTATTTGTGACAGTATGCAAGCTCATCTCTCTCTCT 449
 QY 343 TCGTGTGGGCTTTTACCGTGAAGCGCTGCGTGGAGACCCGTTGGTGAACCAATGACGAAC 402
 Db 450 TCGTGTGGCTTTTATGTGAACCTGTGGTGAACCGTGTGGAGCGACAGTACTTATGCA 509
 QY 403 TGGCGTGGCGGACCGGCTCATGAGCCTGTGTGGGCTTGTGTGAAGGACAAGGACGAGC 462
 Db 510 TGGCCTGTCCGAGCGGCTCATGTGCGTGTGGGGGACCGGTGCACGAGCGGACGAGC 569
 QY 463 AAGCGGAGCTCTGGGGGACAGCTCATTCGCTACGCGCAACTGGGCAAGTGTCTATCC 522
 Db 570 GCGGCGGCGCTTACCGGGGACACTCATGTGCGTACGAGGGCTCTGGCGGTCTATCC 629
 QY 523 TGGCAGAGGTGAGCAGCGCAGTCTACAAGGGCTTCCCAAGCGCCACAGCATGTGTCAAG 582
 Db 630 TGGCGTCCGTGAGCACCGCGGTGTCAAGGGCTTCCCAACATAGACCAAGTGTGGAGG 689
 QY 583 CAGGCTTTATGACTCCGGCAGAAACAAGAGATTGGAAGAACTGAGCTTACACACAACA 642
 Db 690 CTGGGTTTATGACCCCGGAGAGACGCAAGAAAGTTTAAAACTGAACTCATCTTACAAACA 749
 QY 643 TGTTCGGGCTGCGCTGGAGTGTGTTTGGCAACTGTTCATGAAAGGCGTGGCTTGAAGTTC 702
 Db 750 AGTACTGGGTGCGCTGGTGTGTGTTCTCAACTGGCGGCAAGGCGCCACCGAGGGGCG 809
 QY 703 GAATCCGGAGCCCTATTCGTCTCAAGGCTGTGTAACGATGAACAACCTTGTGGTACTTC 762
 Db 810 GCATCCGACCAACAGGCCCTTAACTGTCTCTGAGAGAGCTGATATGTTTTTCCGGGCA 869

Qy 763 AGTGTGACACCTGTATGCTTACGACGTGATTAGTATCCACTGTGTATACACAGGTGG 822
Db 870 AATGTGAATGCTCTTTCATCTATGACTGATTAGCGTACCCTCGTATACAGCGAGGTGG 929
Qy 823 TGACTGTGGCGGTGTACAGCTTCTTCTGACTTGTCTAGTTGGGGGCAATTTCTGAACC 882
Db 930 TGACCATGTGCACCTGTACAGCTACTTCTGCTGCTCATTTGGTGGCCAGTTCCCTGGACC 989
Qy 883 CAGCCAAAGGCTTACCTGCGGCAATGAGCTGACCTGTTGGCCGCTTTCACGTTCTGCG 942
Db 990 CGGCTCAGGGTTACAAAGACACGACCTTAGACCTGTGTGTGCCATCTTCACCTCTTGGC 1049
Qy 943 AGTTCTTCTCTATGTGTGCTGCTGAGGTGG 975
Db 1050 AGTTCTTCTTCTACGCGCGCTGCTCAAGGTAG 1082

RESULT 5
US-10-492-032-30
; Sequence 30, Application US/10492032
; Publication No. US20050049212A1
; GENERAL INFORMATION:
; APPLICANT: Steuermagel, Arnd
; APPLICANT: Brommer, Gunter
; APPLICANT: Fritsch, Rudiger
; APPLICANT: Eulenberger, Karsten
; APPLICANT: Closssek, Thomas
; TITLE OF INVENTION: Beastrophin and Beastrophin Homologous Proteins Involved in the
; TITLE OF INVENTION: Regulation of Energy Homeostasis
; FILE REFERENCE: 2923-606
; CURRENT APPLICATION NUMBER: US/10/492,032
; PRIOR FILING DATE: 2004-04-08
; PRIOR APPLICATION NUMBER: PCT/EP02/11321
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: EP01124059.5
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 30
; LENGTH: 1530
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-492-032-30

Query Match 17.7%; Score 431; DB 19; Length 1530;
Best Local Similarity 68.4%; Pred. No. 9e-112;
Matches 596; Conservative 0; Mismatches 275; Indels 0; Gaps 0;

Qy 105 ATGACCATCACTTACACAGCAAGTGTGCTATATGCGCCCTTAGGCTCTTCTCCCGCTCG 164
Db 1 ATGACCGTACCTACACAGCCGAGTGTGCGAACCGCCGCTTCGGTGGCTTCTCCCAAGCTG 60
Qy 165 CTGCTGTGCTGGCGGGGAGCATCTACAAAGCTCTATATGCGAGTTCTTATCTTCTCG 224
Db 61 CTGCTACTGTGGCGGGAGCATCTACAACTCCTGTGGCAGAGCTCTCGCTTCTCTT 120
Qy 225 CTCGTCTACTACATATCCGCTTATTTATAGAGCTGGCCCTTCAGGAGAAACAAGCTG 284
Db 121 GGGTTCTACATGAGCGGTGAGTGTGCTTACCGCTTTTGTGTGACCGAAGGGCAGAGGCG 180
Qy 285 ATGTTTGAAGAAGTACTCTGATGTGCGACAGCTTACATCCAGCTCATCCCATTTCTTC 344
Db 181 TACTTCGAGAGCTTGTGATTTATTTGTGACAGATGCGAGCTCATCTCTCTCTTC 240
Qy 345 GTGCTGGGCTTCTACGTGAGCTGTGCTGTGACCCGCTGGTGGAAACAAGTACGAGAACTG 404
Db 241 GTGCTTGGCTTTTATGTGACCGCTGGTGTGAACCGCTGGTGGACAGTACTATGATCATG 300
Qy 405 CCGTGGCCGCAACCGCTCATAGAGCTGTGTCTGGGCTTCGTGCAAGGAGAAAGAGAGAA 464
Db 301 CCGCTGGCCGCAACCGCTCATGTGCTGTGTGGGGGCAACGTGCAAGGAGGAGAGACGCG 360
Qy 465 GGGCGGCTGTGCTGGCGCACGCTCATCGCTTACCGCAACCTGGGCAAGTGTCTCATCTCG 524

Db 361 GGGCGGCTCTACCGGCGCACACTATGCGCTTACCGAGGGCTCTCGCGCGCTCATCTCG 420
Qy 525 CGAGCGTACACACCGCAGCTTACAAAGGCTTCCCGAGGCGCCAGCAGCTGTGAAACA 584
Db 421 CGCTCCGTACACCGCGGCTGTTCAGGCTTCCCGACCATAGACACAGTGTGAGGCT 480
Qy 585 GCGTTTATGACTCCGGCGAACAACAAGAGTGTGAGAAATAGAGCTTACCAACAATG 644
Db 481 GGGTTTATGACCCCGAGAGGAGCGCAAGAGTTTGAACCTGAACTCATCTTACAAACAAG 540
Qy 645 TTTGGGTTGCGCTGGGTGTGTTGTCACCTGTTCATGAAAGGCGCTTGGAGTGA 704
Db 541 TACTGGGTGCGCTGGGTGTGTTCTTCCAACTGGGCGGACAGCGCCGAGAGGCGCG 600
Qy 705 ATCCGGACCTTATCTGCTCCAGAGCTGTGTAACGAGATGAACACTTGTGCTACTAG 764
Db 601 ATCCGGACCAACACCGCCTTACCTGTGCTCGAGAGCTGAATGTTTTCGGGCGCAA 660
Qy 765 TGTGACACCTGTATGCTTACGACTAGATTAGTATCCCACTGTGTATPACAGAGTGTG 824
Db 661 TGTGAATGCTCTTCTTCACTATGACTGATTAAGCTTACCTCTGTGTACACGAGTGTG 720
Qy 825 ACTGTGGCGGTGTACAGCTTCTTCTGTGACTTGTATGTGGCGGCACTTCTGAACCA 884
Db 721 ACATCGCAGTGTACAGCTACTTCTGTGCTTGTGCTCATTTGTGCGCACTTCTGAGCCG 780
Qy 885 GCCAAGGCTTACCTGTGCGATGAGTGAACCTGTGTGGCCGCTTCTTACAGTCTGAG 944
Db 781 GCTCAGGGTTTACAAAGACACAGACCTTGAACCTGTGTGTGCCATCTTCAACCTCTTGGAG 840
Qy 945 TTTCTTCTTATGTGTGCTGTGCTGAGGTGG 975
Db 841 TTTCTTCTTCTACGCGCGCTGCTCAAGGTAG 871

RESULT 6
US-10-071-766-10/C
; Sequence 10, Application US/10071766
; Publication No. US20020192678A1
; GENERAL INFORMATION:
; APPLICANT: Huel-Mei Chen
; TITLE OF INVENTION: GENES EXPRESSED IN SENESENCE
; FILE REFERENCE: PA-0043 US
; CURRENT APPLICATION NUMBER: US/10/071,766
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 144
; SOFTWARE: PERL Program
; SEQ ID NO 10
; LENGTH: 1326
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incycle ID No. US20020192678A1 1138151.2
US-10-071-766-10

Query Match 14.4%; Score 350.4; DB 13; Length 1326;
Best Local Similarity 99.7%; Pred. No. 8.2e-89;
Matches 351; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2053 AGCAATTTCTTAACTGCTTCTTAAATGGGAGTGTGGCCAGCGAGGCTTCACTGTGT 2112
Db 1326 AGCAATTTCTTAACTGCTTCTTAAATGGGAGTGTGGCCAGCGAGGCTTCACTGTGT 1267
Qy 2113 GTACACGACGAGACACTGTATCCAGTCAAGCCATATAGCTGTCCACACTGAAAGCTG 2172
Db 1266 GTACACGACGAGACACTGTATCCAGTCAAGCCATATAGCTGTCCACACTGAAAGCTG 1207
Qy 2173 TCTTACCAACGCTTGAATCAATATGTTAGCTTATATGATTAATAATCCAGACTACTGAG 2232
Db 1206 TCTTACCAACGCTTGAATCAATATGTTAGCTTATATGATTAATAATAATCCAGACTACTGAG 1147

QY 2233 CTTTAAATGCTTTTATTCATAAAACGTGGAAGCTAGACATGGAACATTTGAAACATTT 2232
Db 1146 CTTTAAATGCTTTTATTCATAAAACGTGGAAGCTAGACATGGAACATTTGAAACATTT 1087
QY 2293 AACTGACACTCTTGATTCAGAGTCGGGAAACCTTATGTTCTATCTGAATCCAAAGACGCCA 2352
Db 1086 AACTGACACTCTTGATTCAGAGTCGGGAAACCTTATGTTCTATCTGAATCCAAAGACGCCA 1027
QY 2353 CACTTATGATATCTGCGCCAAACATTAATGAGTTTAAATATACAAATCTCGTT 2404
Db 1026 CACTTATGATATCTGCGCCAAACATTAATGAGTTTAAATATACAAATCTCGTT 975

RESULT 7

US-10-492-032-7
Sequence 7, Application US/10492032
Publication No. US20050049212A1
GENERAL INFORMATION:
APPLICANT: Steuernagel, Arnd
APPLICANT: Bronner, Gunter
APPLICANT: Fritsch, Rudiger
APPLICANT: Eulenberg, Karsten
APPLICANT: Ciosek, Thomas
TITLE OF INVENTION: Bestrophin and Bestrophin Homologous Proteins Involved in the
FILE REFERENCE: 2923-606
CURRENT FILING DATE: 2004-04-08
CURRENT APPLICATION NUMBER: US/10/492,032
PCT/EP02/11321
PRIOR FILING DATE: 2002-10-09
PRIOR APPLICATION NUMBER: EP01124059.5
PRIOR FILING DATE: 2001-10-09
NUMBER OF SEQ ID NOS: 37
SOFTWARE: PatentIn version 3.2
SEQ ID NO 7
LENGTH: 1422
TYPE: DNA
ORGANISM: Homo sapiens
US-10-492-032-7

Query Match

Best Local Similarity 14.0%; Score 339.4; DB 19; Length 1422;
Matches 505; Conservative 0; Mismatches 276; Indels 0; Gaps 0;

QY 105 ATGACCATCACTTACCAAGCAAGTGGCTAATGCGCTTATGCTCTTCTCCGCTG 164
Db 1 ATGACCGTTTCATATACCTCTCAAGTGGCGAGCGCCCTTCGAGGTTTCTCGGCTG 60
QY 165 CTGCTGCTGGCGGGGAGCATCTACAAAGCTGTATATGCGAGTTCTTAACTTCTG 224
Db 61 CTTCCTCGGTGAGGAGGAGCATCTACAAAGCTCTCTCAAGGAATCTCTCTTTGGG 120
QY 225 CTCTGCTACTACATCATCCGCTTATTTATAGGCTGGCGCTTACAGGAAGCAAGCTG 284
Db 121 GCCTTGACGTGCTGCTTACATCAACGCGCTGCTGACCCAGGAGAGGTAC 180
QY 285 ATGTTGAGAACTGACTCTGTATTTGACAGACTACATCCAGCTCATCCCATTTCTTC 344
Db 181 GGTATGCTCAGTGGCCCGGTACTGCAACCGCTCAAGACCTCATCTCTTCTTCTT 240
QY 345 GTGCTGGGCTTCTAGTGAAGCTGTGTGACCCGCTGTGGAAACAGTACGAACTG 404
Db 241 GTATTTGGTTTCTATGTAATCTCTGTGTGAACCGCTGTGTGTCACAGTACAAAGCATC 300
QY 405 CCGTGGCCGACCGCTCATGAGCTGTGTGCGGCTTGTGCAAGGAAGCAAGCAAGCA 464
Db 301 CCGCTGCGACAGCTGATGTGTGCTCATCTCGGTAGGTGACGCGTGGACAGCGG 360
QY 465 GGCAGGCTGCTGCGGCGGACGCTCATCCGCTTACGCAACCTGTGGCAACGTTCTCATCTG 524
Db 361 GGCAGGCTGCTGCGGCGGACGCTCATCCGCTTACGCAACCTGTGGCAACGTTCTCATCTG 420
QY 525 CGAGGCTGACACCGGAGTCTACAGCGCTTCCCAAGGCGCCAGACACTGTGTGACAGA 584

Db 421 CGCTCGTACAGACCCGCGTCTTAAAGCGCTTCCCAACATGAGGACGTGTGAGCGCA 480
QY 585 GCGTTTATGACTCCGCGCAACACAGAGCTTGGAGAACTGAGCCTACCAACAATG 644
Db 481 GGTTCATGTCCCAAGAGAGAGGAAAGTTTGAAGCCCTGAATCCGACTTCAACAG 540
QY 645 TTCTGGGAGCCCTGGGAGTGTGTTTGGCAACTGTCAATGAAGCGTGTGGAGTGA 704
Db 541 TACTGGGTCCTGCGTGTGTTTCAACCACTGGCGGCCAGGCGCCGAGAGCGGCGCA 600
QY 705 ATCCGGAACCTATCTGCTCCAGAGCTGTGCAAGAGATGAACAACCTTGCGTACAG 764
Db 601 ATACGTGACGATATGCTCTCTGTCTATTTGGAAAGCTGAACAGTACCAAGCGCAAG 660
QY 765 TGTGAGACCTGTATGCTGCTAGATGATTAATGATCCACTGTGTATACAGTGTG 824
Db 661 TGACAGATGCTATTCATCCATGACTGATGAGATCCCTCTGCTTACACCCAAAGTGTG 720
QY 825 ACTGTGGCGGTATACAGCTTCTTCTGACTGTGTGTGAGTGGCGGAGTTTGAACCA 884
Db 721 ACCATAGCCGCTACTCTTCTTGGCCCTCTCCCTGTGTGGCGGAGTTTGAAGCCA 780
QY 885 G 885
Db 781 G 781

RESULT 8

US-10-108-260A-1299
Sequence 1299, Application US/10108260A
Publication No. US20040005560A1
GENERAL INFORMATION:
APPLICANT: HELIX RESEARCH INSTITUTE
TITLE OF INVENTION: NO. US20040005560A1 full length cDNA
FILE REFERENCE: H1-A0106
CURRENT APPLICATION NUMBER: US/10/108,260A
CURRENT FILING DATE: 2002-03-27
NUMBER OF SEQ ID NOS: 5458
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1299
LENGTH: 2500
TYPE: DNA
ORGANISM: Homo sapiens
US-10-108-260A-1299

Query Match

Best Local Similarity 13.9%; Score 338.2; DB 17; Length 2500;
Matches 538; Conservative 0; Mismatches 333; Indels 0; Gaps 0;

QY 105 ATGACCATCACTTACCAAGCAAGTGGCTAATGCGCTTATGCTCTTCTCCGCTG 164
Db 229 ATGACGTGCTACTTACTCAAGTAAGTAGCAATGCACTTTTGGATTTCAATAGTTA 288
QY 165 CTGCTGCTGGCGGGGAGCATCTACAAAGCTGTATATGCGAGTTCTTAACTTCTG 224
Db 289 CTCTCAAGTGGAGAGGACATCTACAAAGCTGTATGAGGAATTTATGTTTGTCT 348
QY 225 CTCTGCTACTACATCAACCGCTTATTTATAGGCTGGCGCTTACGGAAGCAAGCTG 284
Db 349 GTTCTTTATACAGCAATAGTTGTGTGACAGATTTTATCAAGAGGTCCAAAACGT 408
QY 285 ATGTTGAGAACTGACTGTATTTGACAGACTACATCCAGCTCATCCCATTTCTTC 344
Db 409 TACTTTGAAAAATTATCAATTTATCTGTACATATGCTGAACAAATTCAGTAACCTTT 468
QY 345 GTGCTGGGCTTCTAGTGAAGCTGTGTGACCCGCTGTGGAAACAGTACGAACTG 404
Db 469 GTGCTGGGCTTATATGTTACTCTGTGTGAACCGATGTGTGAACCAAGTTTGTGAATTTG 528
QY 405 CCGTGGCCGACCGCTCATGAGCTGTGTGCGGCTTGTGCAAGGCAAGAGCAAGCA 464
Db 529 CCTGGCCAGAGAGGTATATGTTCTCTCATCTTACAGATTTTCAAGGAAGCAAGAGCAC 588

Qy	465	GGCGGCGTGC	CGCGCGACGCTATCGGCTACGCGAACCTGGGACAGTGCTATCCTG	524
Db	589	GGGGCGCCCTG	CTTAAAGGACGCTGATGGCTACGCTCAATCTCACTCCCTGCTCATCTTT	648
Qy	525	CGCAGCGCTAC	ACCGCAGTCTACAAAGCGCTTCCCGACGGCCGACGACCTGGTCAAGCA	584
Db	649	CGCTCGGAGACA	CTGCTGTGTGTAACAAAAGATTTCCCAATGACACAGTGCTGGAGCA	708
Qy	585	GGCTTTATGACT	CCGCGGAGAAACAAGCAGTTGGAGAACTGAGCTTACACACAACATG	644
Db	709	GGTTTATGACA	ACAGATGAAGAAATTAATTCACACACTCAAGTCTCTCATCTGAAA	768
Qy	645	TTCTGGGTG	CGCCCTGGGTGTGGTTGGCCAACTGTATGAAAGCGGTGGCTTGGAGTGTGA	704
Db	769	TATTGGGTTCAT	TCTCATCTGGTTGGAAATCTTGGACATTAAGCCGGGAAATGAAGGTGAA	828
Qy	705	ATCCGGGACCC	ATATCCGCTCAGAGCGCTGCGTGAACGAGATGAACACTTGGGTACTCAG	764
Db	829	ATCAGAGCAG	GTGTTGATCTGTAATCATTTGATGATGAAATGAATTCGATACCGCTCTTGG	888
Qy	765	TGTGGACACT	GTATGCTTACGACTGATTAATGCCACTGTGTATACAGAGTGGTGTG	824
Db	889	TGCAGCCCTTA	TTCGTTATGATGACTGGGTGGGATTCGCTGGTTTACACCCAGGTGTC	948
Qy	825	ACTGTGGCGGT	ATACAGCTTCTTCTCTGACTTGTCTAGTGTGGCGGACGATTTCTGAACCA	884
Db	949	ACTCTGTGTG	TCTATACCTTCTTCTTTCGATGCGCTGATTTGACGCGACAGTTTGTGATCC	1008
Qy	885	GCCAAAGGCTA	ACCTGGCAGTGAAGTGAACCTCGTTGTGCCCGTCTTACGTTCTCTGAC	944
Db	1009	ACCAAAGGCTA	CGCAGGGCAGTACTGTGATCTTTAATTCCTCATCTTCAACCTCCTCTCAAA	1068
Qy	945	TTCTTCTTCTA	TGTTGGCTGAGGTGAGGTGG	975
Db	1069	TTCTTCTTCTA	TGAGATGGCTTTAAGGTAG	1099

```

RESULT 9
US-10-492-032-5
? Sequence 5, Application US/10492032
? Publication No. US20050049212A1
? GENERAL INFORMATION:
? APPLICANT: Steuernagel, Arnd
? APPLICANT: Bronner, Gunter
? APPLICANT: Pritsch, Rudiger
? APPLICANT: Eulenberg, Karsten
? APPLICANT: Closssek, Thomas
? TITLE OF INVENTION: Bestrophin and Bestrophin Homologous Proteins Involved in the
? FILE REFERENCE: 2923-606
? CURRENT APPLICATION NUMBER: US/10/492,032
? CURRENT FILING DATE: 2004-04-08
? PRIOR APPLICATION NUMBER: PCT/EP02/11321
? PRIOR FILING DATE: 2002-10-09
? PRIOR APPLICATION NUMBER: EP01124059.5
? PRIOR FILING DATE: 2001-10-09
? NUMBER OF SEQ ID NOS: 37
? SOFTWARE: PatentIn version 3.2
? SEQ ID NO 5
? LENGTH: 2028
? TYPE: DNA
? ORGANISM: Homo sapiens
US-10-492-032-5

Query Match      13.9%; Score 336.8; DB 19; Length 2028;
Best Local Similarity 61.8%; Pred. No. 7.7e-85;
Matches 536; Conservative 0; Mismatches 332; Indels 0; Gaps 0;

105 ATGACATCATCTTACACAAAGCCAAAGTGCTATGCGCGGTTAGGCTCCTTCTCCGCGCTG 164
|||||
1 ATGATGTCATCACTACTCTCCAGTAAAGTAGCAATGCAACTTTTGGATTTGCATAGGTTA 60

```

QY	165	TTGCTGTGCTCGGGGGGAGCATCTACAAAGCTGTATATGGGAGTCTTAATCTTCCTG	224
Db	61	CTCTCAAGTAGAGGGAGCATCTACAAACTATGTACAGGAAATTAATGTGTTTTGCT	120
QY	225	CTTGCTACTACATATATCCGCTTTATTTATATAGCTGGCCCTCAGCGGAACAACAGCTG	284
Db	121	GTTCCTTTATACAGCAATAAGTTTGGTATATACAGATTGTTAATTACAGGAACTCCAAAACGT	180
QY	285	ATGTTTGAAGAACTGACTCTGTATTTGCGACAGCTACATCCAGCTCATCCCCATTTCTTC	344
Db	181	TACTTTGAAAAATTAATCAATTTACTGTGACAGATATGTGAAACAAATTCAGATTAACCTTT	240
QY	345	GNCTGGGCTCTACGTACGCTGGTTCGNAAGCCGCTGGTGAACACAGTACAGAACTTG	404
Db	241	GTGCTTGGGTTTATGTTTACTCTGTGTATGAAACCAATGTGTGAACCAAGTTGTGAATTTG	300
QY	405	CCGTGGCCGACCCGCTCATGAGCCCTGTGTGCGGCTTTGCTGAAAGCAAGACGAGCAA	464
Db	301	CCCTGGCCAGACAGGCTAATGTTCTCATCTTACAGAGTTTACAGGAAGCGACGAGCAC	360
QY	465	GGCGGCTGCTGCGGGCGCAGCTCATCCGCTACGCGACACTTGGGCAACGTCTCATCTCG	524
Db	361	GGGCGCCCTGCTTGAAGAGACGCTGATGGCTACGTCATCTCACTCCCTGCTCATCTTT	420
QY	525	CGACGCTCAGACACGCAATCTACAAAGCGCTTCCCCAGCGCCAGACACTGGTGCAGACA	584
Db	421	CGCTGGGTGAGCATCTGCTGTGTACAAAAGATTTCCCAATATGACCAAGTGGTGAAGCA	480
QY	585	GAGCTTATGACTCCCGGCAACACACAGCAGTTGAGAAACTGAGCCTTACACACACATG	644
Db	481	GGTTTATATGACAAACAGATGAAGAAATTAATTAACACACTCAAGTCTCATCTGAA	540
QY	645	TTCTGGGTGCCCTGGGTGTGTTTGTCCCAACCTGTCAATGAAAGCGTGGCTTGAAGTCA	704
Db	541	TATTTGGTTCATTTATCTGTGTTTGGAAATCTTGCACCTAAAGCCCGGAATGAAGGTAGA	600
QY	705	ATCCGGAGCCCATATCTGCTCCAGAGCGCTGCGAAAGATGAACACTTTGGTACTCAG	764
Db	601	ATCAGAGACAGTTGATCTGTGAAATCATTTGATGACTGAATGAATGATACGGCTTTGG	660
QY	765	TGTGGACACTGTATGCTTACGACTGTGATTAGTATCCACTGTGTATACACAGTGGTGTG	824
Db	661	TGCAGCCCTTAATTCGGTTATGACGTGGTGGGATTTCCGCTGGTTTACACCAAGGTGTCTC	720
QY	825	ACTGNGGGGTGTACAGCTTCTTCTCTGACTTGTCTAAGTTGGGGCGGAGTTTTCGAACCA	884
Db	721	ACTCTGTGCTATACCTTCTCTTTGTGGTCCGTGATTTGGAACGCCAGTTTGTGATCCC	780
QY	885	GCCAAAGGCTTACCCCTGGGCATGAGCTGTGACCTGTGTGCCCCGTCTTCAAGTTTCCGAC	944
Db	781	ACCAAAAGGCTACGCAAGGACATGACTTTGGATCTTTAATTCATTCATCTTCAACCTCCTACAA	840
QY	945	TTCTTCTTCTATGTTGGCTGGCTGAAG	972
Db	841	TTCTTCTTCTATGACAGATGGCTTAAAG	868

RESULT 10
US-09-814-353-19504/C
Sequence 19504, Application US/09814353
Publication No. US20030165831A1
GENERAL INFORMATION:
APPLICANT: Lee, John
APPLICANT: Thompson, Pamela
APPLICANT: Lillie, James
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
FILE REFERENCE: MRI-006B
CURRENT APPLICATION NUMBER: US/09/814,353
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: US 60/191,031
PRIOR FILING DATE: 2000-03-21


```

; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19504
; LENGTH: 620
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-19504
```

```

Query Match      8.8%; Score 213; DB 10; Length 620;
Best Local Similarity 66.4%; Pred. No. 8,6e-50;
Matches 306; Conservative 0; Mismatches 155; Indels 0; Gaps 0;
```

```

QY 103 CCATGACCATCACTTACACAGCCCAAGTGGCTAATGCGCTTAGGCTCTCCGCC 162
DB 464 CCATGACGGTTTCATACACTCTCAAGCGGAGGCCCGCTTGGAGTTTCTTGACC 405
QY 163 TGCCTGCTGCTGGGGGACACATCTACAGCTGCTATGCGAGTTCTTAATCTCC 222
DB 404 TCTCTTCGGCTGAGGGAGACATCTACAGCTCTCTACAGGAATTCCTCTCTTGG 345
QY 223 TGCCTGCTACTACATCACTCCGCTTATTTATTAAGGCTGACCGCTTACGGAAGCAAGC 282
DB 344 GGGCGCTTAGCGCTGCTGCTAGCATCACTACCGGCTGCTGTAACCGAGAGAGGT 285
QY 283 TGATGTTGAGAACTGACTCTGTAATGCGACAGCTACATCCAGCTCAATTCCT 342
DB 284 AGCTATAGCTCAGAGTGCGCGCTGACTGCAACCGCTGACGAGACCTGATTCCTTGTCT 225
QY 343 TCGTCTGAGGCTTCTACGAGCGCTGCTGCTGACCGGCTGCTGGAACGATACGAAACC 402
DB 224 TTGATTTAGGTTTCTATGTAAGCTCTGCTGTAACCGCTGCTGATCCAGTACCAAGCA 165
QY 403 TCCCGTGACCCGACCGCTCTATGAGCTGCTGCGGCTTGTGGAAGCAAGAGCAGC 462
DB 164 TCCGCTGTCAGACCAAGCTGATGCGTCACTGCGCTAGAGTGACCGGCTGAGACAGC 105
QY 463 AAGCGCGCTGCTGCGGAGCAGCTCATCGCTACGCCCACTGGGCAACGCTGCTCATCC 522
DB 104 GGGGCGGCTGCTGCGCGCAGCCCTCATCCGCTACGGCAACTGGGCTTGGTGTGTGTC 45
QY 523 TGCCGACGCTCAGCAGCGCAGTCTTACAGAGGCTTCCCGCAGC 563
DB 44 TGCCTCGGTACAGACCGCGCTGCTTAAGCGCTTCCCGACC 4
```

```

RESULT 11
US-10-198-846-11070
; Sequence 11070, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
```

```

; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11070
; LENGTH: 1292
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1287, 1288, 1289, 1290, 1291, 1292
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-11070
```

```

Query Match      7.8%; Score 189.2; DB 14; Length 1292;
Best Local Similarity 62.4%; Pred. No. 7.7e-43;
Matches 296; Conservative 0; Mismatches 178; Indels 0; Gaps 0;
```

```

QY 349 TGGGCTTCTAGTACCGCTGCTGTAACCCGCTGCTGGAACCGATACGAACTTGGCT 408
DB 318 TAGGTTTATGTAAGTCTGTAAGTGAACCGATGCTGGAACCGATTGTGAATTTGCCCT 377
QY 409 GGGCCGACCGCTCATGAGCGCTGCTGCTGCGGCTTCTGTCGAAGGCAAGACGAGAGGCC 468
DB 378 GGCAGACAGGCTATATTTCTCATCTCTAGCAGTGTTCAGAGCGAGACGAGCAGGGC 437
QY 469 GCGTCTGCGCGCGCAGGCTCATCCGCTAGCGCAACTGGAGCAAGTCTCATCTCGCA 528
DB 438 GCTGCTTAAAGAGCGCTGATGCGCTAGCTCATCTCACTTCCTGCTCATCTTTCGCT 497
QY 529 GCGTACAGCAGCGAGTCTTCAAGCGCTCCCGCGCGCCAGACCGTGGAGCAAGCGCT 588
DB 498 CGGTAGCACTGCTGTGTACAAAAGATTTCACAAAGACAGCGTGTGAAGCAGGTT 557
QY 589 TTATGACTCGGCAAGAACAGACAGCTTGAAGAACTGAGCTTACCCACCAACATGTTCT 648
DB 558 TTATGACAAAGATGAAGAAATTAATTCACACACTCAAGTCTCTCATCTGAATATT 617
QY 649 GGGTGGCGGCTGCTGCTTGGCACTGCTCAATGAAGCGCTGCTGAGAGTCCGATCC 708
DB 618 GGGTTCATTCATCTGTTGTAATCTTGCAACTTAAACCGGGAATGAAGTATGATCA 677
QY 709 GAGACCTTATCTGCTCCAGACCTGCTGAAGAGATGAACCTTGCCTACTCAGTGTG 768
DB 678 GAGACAGTGTATCTGCAATCATTTATGACTGAATGATGATGATACCGCTCTTGTGCA 737
QY 769 GACACTGTATGCTTACGACTGATTAATGATCCACTGCTGATACACAGGTGG 822
DB 738 GCCTTATTCGATTATGACTGGGTTGGGATTCGCTGCTTAAACCCAGGTAG 791
```

```

RESULT 12
US-09-880-107-2174/C
; Sequence 2174, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2174
; LENGTH: 1198
```

```
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Genbank Accession No. US20020142981A1 L20941
US-09-860-107-2174
```

```
Query Match      7.2%; Score 175; DB 9; Length 1198;
Best Local Similarity 100.0%; Pred. No. 8.3e-39;
Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 2230 CAGCCTTTATGCTTTTATTCATTAATAAAGCTGTAAGCTGTAACCAATGGAAACA 2289
DB 1198 CAGCCTTTATGCTTTTATTCATTAATAAAGCTGTAAGCTGTAACCAATGGAAACA 1139
QY 2290 TTTAACTAGACTGCTGATTAAGTGGGAACCTTGTCTATCTGAATCCAGACAG 2349
DB 1138 TTTAACTAGACTGCTGATTAAGTGGGAACCTTGTCTATCTGAATCCAGACAG 1079
QY 2350 CCACACCTTAGTATTAAGTCCCAACTAATGATTAATTAATCAATAATCTGTT 2404
DB 1078 CCACACCTTAGTATTAAGTCCCAACTAATGATTAATTAATCAATAATCTGTT 1024
```

```
RESULT 13
US-10-027-632-134530/c
Sequence 134530, Application US/10027632
Publication No. US20020198371A1
GENERAL INFORMATION:
```

```
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: PasteSeq for Windows Version 4.0
SEQ ID NO 134530
LENGTH: 578
TYPE: DNA
ORGANISM: Human
US-10-027-632-134530
```

```
Query Match      6.6%; Score 161.4; DB 13; Length 578;
Best Local Similarity 99.4%; Pred. No. 4.3e-35;
Matches 162; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 582 GCAGGCTTTATGACTCCGGGAGAACACAGAGATTGGAAAGTGAAGCTTCCACACAC 641
DB 183 GCAGGCTTTATGACTCCGGGAGAACACAGAGATTGGAAAGTGAAGCTTCCACACAC 124
QY 642 ATGTTCTGGGTGCTGCTGGGTGCTTTGCCAAGCTGTCATGAAAGGCGTGGCTGAGAGT 701
DB 123 ATGTTCTGGGTGCTGCTGGGTGCTTTGCCAAGCTGTCATGAAAGGCGTGGCTGAGAGT 64
QY 702 CGAATCCGGGACCTATCTGCTCCAGAGCTGCTGTAACGAGA 744
DB 63 CGAATCCGGGACCTATCTGCTCCAGAGCTGCTGTAACGAGA 21
```

```
RESULT 14
US-10-027-632-134530/c
Sequence 134530, Application US/10027632
Publication No. US20030204075A9
GENERAL INFORMATION:
```

```
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: PasteSeq for Windows Version 4.0
SEQ ID NO 134530
LENGTH: 578
TYPE: DNA
ORGANISM: Human
US-10-027-632-134530
```

```
Query Match      6.6%; Score 161.4; DB 17; Length 578;
Best Local Similarity 99.4%; Pred. No. 4.3e-35;
Matches 162; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 582 GCAGGCTTTATGACTCCGGGAGAACACAGAGATTGGAAAGTGAAGCTTCCACACAC 641
DB 183 GCAGGCTTTATGACTCCGGGAGAACACAGAGATTGGAAAGTGAAGCTTCCACACAC 124
QY 642 ATGTTCTGGGTGCTGCTGGGTGCTTTGCCAAGCTGTCATGAAAGGCGTGGCTGAGAGT 701
DB 123 ATGTTCTGGGTGCTGCTGGGTGCTTTGCCAAGCTGTCATGAAAGGCGTGGCTGAGAGT 64
QY 702 CGAATCCGGGACCTATCTGCTCCAGAGCTGCTGTAACGAGA 744
DB 63 CGAATCCGGGACCTATCTGCTCCAGAGCTGCTGTAACGAGA 21
```

```
RESULT 15
```

```
US-10-027-632-134529/c
Sequence 134529, Application US/10027632
Publication No. US20020198371A1
GENERAL INFORMATION:
```

```
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
```

; PRIOR APPLICATION NUMBER: US 60/146,002
 ; PRIOR FILING DATE: 1999-08-09
 ; NUMBER OF SEQ ID NOS: 325720
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 134529
 ; LENGTH: 748
 ; TYPE: DNA
 ; ORGANISM: Human
 US-10-027-632-134529

Query Match 6.2%; Score 149.4; DB 13; Length 748;
 Best Local Similarity 90.9%; Pred. No. 1.3e-31;
 Matches 170; Conservative 0; Mismatches 16; Indels 1; Gaps 1;

QY	558	CCCAGCGCCCGACACCTGTGTGCAAGCAGGCTTATGACTCCGCGAGACACAAGCAGTTG	617
Db	712	CCCACCGCTTTTTCACCTCCACTCTGCGAGGCTTTATGACTCCGCGAGACACAAGCAG-TG	654
QY	618	GAGAACTGAGCTTACCAACAACATGTTCTGGGTGCCCTGGGTGTGTTGCCAACCTG	677
Db	653	GAGAACTGAGCTTACCAACAACATGTTCTGGGTGCCCTGGGTGTGTTGCCAACCTG	594
QY	678	TCAATGAGGCGGTGGCTTGGAGGTGGAATCCGGGACCTTATCTGCTCCAGAGCTGCTG	737
Db	593	TCAATGAGGCGGTGGCTTGGAGGTGGAATCCGGGACCTTATCTGCTCCAGAGCTGCTG	534
QY	738	AACGAGA	744
Db	533	AACGTGA	527

Search completed: March 26, 2005, 22:07:42
 Job time : 1362.06 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using bw model

Run on: March 25, 2005, 22:47:44 ; Search time 387.221 seconds
(without alignments)
10264.201 Million cell updates/sec

Title: US-09-622-964a-4
Perfect score: 2429
Sequence: 1 cagggagcaccaccagccta.....aaaaaaaaaaaaaaaa 2429

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PTCUS.COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1954.8	80.5	2171	4 US-09-949-016-4901	Sequence 4901, Ap
2	642.2	26.4	7609	4 US-09-949-016-16644	Sequence 16644, A
3	642.2	26.4	18075	4 US-09-949-016-16643	Sequence 16643, A
4	358.4	14.8	1462	4 US-09-949-016-4902	Sequence 4902, Ap
5	188.4	7.8	601	4 US-09-949-016-173056	Sequence 173056, A
6	161.4	6.6	601	4 US-09-949-016-173051	Sequence 173051, A
7	161.4	6.6	601	4 US-09-949-016-173052	Sequence 173052, A
8	128	5.3	601	4 US-09-949-016-173045	Sequence 173045, A
9	89	3.7	4916	4 US-09-949-016-123932	Sequence 123932, A
10	89	3.7	4916	4 US-09-949-016-15234	Sequence 15234, A
11	88.6	3.6	601	4 US-09-949-016-46055	Sequence 46055, A
12	88.6	3.6	43086	4 US-09-949-016-13055	Sequence 13055, A
13	88.4	3.6	678533	4 US-09-949-016-14577	Sequence 14577, A
14	88.4	3.6	10827	4 US-09-949-016-14578	Sequence 14578, A
15	87.8	3.6	10827	4 US-09-949-016-12297	Sequence 12297, A
16	87.8	3.6	11939	4 US-09-949-016-13232	Sequence 13232, A
17	87.8	3.6	17605	4 US-09-949-016-13599	Sequence 13599, A
18	87.8	3.6	25231	4 US-09-949-016-12041	Sequence 12041, A
19	87.8	3.6	25431	4 US-09-949-016-13234	Sequence 13234, A
20	87.8	3.6	50836	4 US-09-949-016-16722	Sequence 16722, A
21	87.4	3.6	27663	4 US-09-949-016-16160	Sequence 16160, A
22	87.4	3.6	28109	4 US-09-949-016-14329	Sequence 14329, A
23	87.4	3.6	31573	4 US-09-949-016-15628	Sequence 15628, A
24	87.4	3.6	31573	4 US-09-949-016-16327	Sequence 16327, A
25	87.4	3.6	35629	4 US-09-949-016-15786	Sequence 15786, A
26	87.2	3.6	275110	4 US-09-949-016-12706	Sequence 12706, A
27	87.2	3.6	275110	4 US-09-949-016-16070	Sequence 16070, A

28	86.6	3.6	601	4 US-09-949-016-121853	Sequence 121853, A
29	86.4	3.6	9608	4 US-09-949-016-16706	Sequence 16706, A
30	86.2	3.5	601	4 US-09-949-016-28769	Sequence 28769, A
31	86.2	3.5	601	4 US-09-949-016-28770	Sequence 28770, A
32	86.2	3.5	601	4 US-09-949-016-52283	Sequence 52283, A
33	86.2	3.5	601	4 US-09-949-016-52285	Sequence 52285, A
34	86.2	3.5	601	4 US-09-949-016-52286	Sequence 52286, A
35	86.2	3.5	601	4 US-09-949-016-52287	Sequence 52287, A
36	86.2	3.5	601	4 US-09-949-016-77763	Sequence 77763, A
37	86.2	3.5	601	4 US-09-949-016-77764	Sequence 77764, A
38	86.2	3.5	601	4 US-09-949-016-117618	Sequence 117618, A
39	86.2	3.5	601	4 US-09-949-016-126371	Sequence 126371, A
40	86.2	3.5	601	4 US-09-949-016-126372	Sequence 126372, A
41	86.2	3.5	601	4 US-09-949-016-126373	Sequence 126373, A
42	86.2	3.5	601	4 US-09-949-016-168504	Sequence 168504, A
43	86.2	3.5	601	4 US-09-949-016-168552	Sequence 168552, A
44	86.2	3.5	601	4 US-09-949-016-168600	Sequence 168600, A
45	86.2	3.5	16236	4 US-09-949-016-16238	Sequence 16238, A

ALIGNMENTS

```
RESULT 1
US-09-949-016-4901
; Sequence 4901, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C0001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; PRIOR FILING DATE: 2000-04-14, 755
; PRIOR APPLICATION NUMBER: 60/241, 768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/223, 498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4901
; LENGTH: 2171
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-4901

Query Match      80.5%; Score 1954.8; DB 4; Length 2171;
Best Local Similarity 91.4%; Pred. No. 0;
Matches 2169; Conservative 0; Mismatches 2; Indels 203; Gaps 1;
```

QY	17	CTTAGTCCGACGACCTTCTGTGGATCATCGGACCCACCTTGAAACCCCACTGACCCAG	76
DB	1	CTTAGTCCGACGACCTTCTGTGGATCATCGGACCCACCTTGAAACCCCACTGACCCAG	60
QY	77	CCCACTGTCGACGACCCACCTGCGGACGACCATCATCTTACCAAGCCAAAGGCGTAA	136
DB	61	CCCACTGTCGACGACCCACCTGCGGACGACCATCATCTTACCAAGCCAAAGGCGTAA	120
QY	137	TGCCCGCTTAGGAGCTTCTTCCGCGCTGCTGTGCTGGCGGGGACGACTCTCAAGCT	196
DB	121	TGCCCGCTTAGGAGCTTCTTCCGCGCTGCTGTGCTGGCGGGGACGACTCTCAAGCT	180
QY	197	GCTATATAGGCGAGTTCTTAATCTTCTGCTCTGCTACTATCATCATCCGCTTATTATAG	256
DB	181	GCTATATAGGCGAGTTCTTAATCTTCTGCTCTGCTACTATCATCATCCGCTTATTATAG	240
QY	257	GCTGGCCCTTACGGAAGAAACAAGCGTGAATTTGAAGAACTGCTGTATTGCGACAG	316
DB	241	GCTGGCCCTTACGGAAGAAACAAGCGTGAATTTGAAGAACTGCTGTATTGCGACAG	300
QY	317	CTACATCCAGCTCATCCCATTTCTCTGCTGGGCTTCTACCTGACGCTGTCTGAC	376

Db 301 CTACATCCAGCTCATCCCATTTCTTCTGCTGAGCTTCTACGAGAGCTGAGCTGAC 360
 Qy 377 CCGCTGTGGAAACAGTAAGAACTGCGCTGAGCCGACCGCCATGAGCCTGATGTC 436
 Db 361 CCGCTGTGGAAACAGTAAGAACTGCGCTGAGCCGACCGCCATGAGCCTGATGTC 420
 Qy 437 GGGCTTCTGTAAGGCAAGAGCAAGCAAGGCGGCTGAGGAGCAAGCTCATGCGCTA 496
 Db 421 GGGCTTCTGTAAGGCAAGAGCAAGCAAGGCGGCTGAGGAGCAAGCTCATGCGCTA 480
 Qy 497 GCGCAACTGAGGCAAGCTGCTCATCTGCGAGCGCTGAGCAAGCAAGCTCATGAGCGCTT 556
 Db 481 GCGCAACTGAGGCAAGCTGCTCATCTGCGAGCGCTGAGCAAGCAAGCTCATGAGCGCTT 540
 Qy 557 CCGCAGCGCCAGCACTGCTGAGCAAGGAGCTTTATGATCCGCAAGCAAGCAAGCTT 616
 Db 541 CCGCAGCGCCAGCACTGCTGAGCAAGGAGCTTTATGATCCGCAAGCAAGCAAGCTT 600
 Qy 617 GGAAGAACTGAGCTTACCAACAAATGTTCTGAGGCTCCGAGGAGTGTGCTGCAAGCTT 676
 Db 601 GGAAGAACTGAGCTTACCAACAAATGTTCTGAGGCTCCGAGGAGTGTGCTGCAAGCTT 660
 Qy 677 GTCAATGAGGCGTGGCTGAGAGTCCGATCCGAGACCTTATCTGCTGCAAGCTT 736
 Db 661 GTCAATGAGGCGTGGCTGAGAGTCCGATCCGAGACCTTATCTGCTGCAAGCTT 720
 Qy 737 GAAAGAGATGAACACTTGGCTACTAGTGTGACACTGTATGCTTATGAGCTGATG 796
 Db 721 GAAAGAGATGAACACTTGGCTACTAGTGTGACACTGTATGCTTATGAGCTGATG 780
 Qy 797 TATCCAGTGTGATACAGAGTGTGATCTGAGGAGTGTGACAGCTTCTGCTGATG 856
 Db 781 TATCCAGTGTGATACAGAGTGTGATCTGAGGAGTGTGACAGCTTCTGCTGATG 840
 Qy 857 TCTAGTGTGGGCGGCACTTCTGAAACCAAGCAAGGCTTACCTGAGCAAGCTGAGCTT 916
 Db 841 TCTAGTGTGGGCGGCACTTCTGAAACCAAGCAAGGCTTACCTGAGCAAGCTGAGCTT 900
 Qy 917 CGTTGTGCGGCTTCTGACCTTCTGCAAGTGTCTTCTGATGTTGAGCTGAGAGTGG 976
 Db 901 CGTTGTGCGGCTTCTGACCTTCTGCAAGTGTCTTCTGATGTTGAGCTGAGAGTGG 954
 Qy 977 CCTCTCAAGGCGCTGCTGAGGCTGAGGCAAGGAGTGTGAGGCTGAGGCTGAGCTT 1036
 Db 955 ----- 954
 Qy 1037 TGAAGAGAGATGCACTGTCAAGAAAGAGGCTTCAAGGAGTGAAGAGCAAGCAGGCTG 1096
 Db 955 ----- 954
 Qy 1097 GTGGCGCACTGTATATCCAGCTTACTCGGAGGCTGAGGAGAGATGCTTGAACC 1156
 Db 955 ----- 954
 Qy 1157 CCGGAGGCGGAGTGTGCTGAGAGAGCTTCAACCCCTTGTGAGAGATGATGAT 1216
 Db 955 ----- 954
 Qy 1217 GATTTTGAACCAATGATTTGAGAGAGATTTGAGAGTGTGCTGTTGCTGAGAT 1276
 Db 998 GATTTTGAACCAATGATTTGAGAGAGATTTGAGAGTGTGCTGTTGCTGAGAT 1057
 Qy 1277 GAGATGCAACAGGACTGCTGAGTGAAGCCGAGCACTGATGAGATGAAGCCGAGCA 1336
 Db 1058 GAGATGCAACAGGACTGCTGAGTGAAGCCGAGCACTGATGAGATGAAGCCGAGCA 1317
 Qy 1337 CAGCCCCCTTACAGAGTGTGCTGAGGCTTCAAGCTTCTGATGAGCTTCAAC 1396
 Db 1118 CAGCCCCCTTACAGAGTGTGCTGAGGCTTCAAGCTTCTGATGAGCTTCAAC 1177
 Qy 1397 TTCAATCAGCTGAAACAAAGAGATGAGTTCAGCCCAATGAGAGAGAGAG 1456

Db 1178 TTCAATCAGCTGAAACAAAGAGATGAGATTTCCAGCCCAATCAGAGAGAGAG 1237
 Qy 1457 GATGCTCAGCGTGCATATTTGGCGCTTCTTGAAGCCCTGAGTCCATGATCATTCT 1516
 Db 1238 GATGCTCAGCGTGCATATTTGGCGCTTCTTGAAGCCCTGAGTCCATGATCATTCT 1297
 Qy 1517 CCGAGGCAAACTCAAGGACAAACTGATGAGGCGCAAGAGGAAATCCCTTCTCAGAG 1576
 Db 1298 CCGAGGCAAACTCAAGGACAAACTGATGAGGCGCAAGAGGAAATCCCTTCTCAGAG 1357
 Qy 1577 GGCCTGCGCAAAACCAAGAGGAGCAAGAGAGCTTGAAGGCGCAGAGCAAGCAAG 1636
 Db 1358 GGCCTGCGCAAAACCAAGAGGAGCAAGAGAGCTTGAAGGCGCAGAGCAAGCAAG 1417
 Qy 1637 GCTTGAAGCTTGAAGGCGTGAAGGCGCTTCAAGTCTGAGGCGCAAGTATCAGAGGCG 1696
 Db 1418 GCTTGAAGCTTGAAGGCGTGAAGGCGCTTCAAGTCTGAGGCGCAAGTATCAGAGGCG 1477
 Qy 1697 TACTACAGTGCCTCCACAGAGCGCCCTCAGCCCACTCCATGTTCTTCCCTGAGAGCA 1756
 Db 1478 TACTACAGTGCCTCCACAGAGCGCCCTCAGCCCACTCCATGTTCTTCCCTGAGAGCA 1537
 Qy 1757 TCAAGCGCTGCAAGCTTCAAGTGTCAAGGCTGAGCAACCAAGCAAGAGCTTAAAG 1816
 Db 1538 TCAAGCGCTGCAAGCTTCAAGTGTCAAGGCTGAGCAACCAAGCAAGAGCTTAAAG 1597
 Qy 1817 ACTGTGAGTCTGAGGCGCAAGAAAGTTTGAATGCTCTGAGAGAGAGAGTGGGCTT 1876
 Db 1598 ACTGTGAGTCTGAGGCGCAAGAAAGTTTGAATGCTCTGAGAGAGAGAGTGGGCTT 1657
 Qy 1877 ATGAGAGCCCAAGAAATATCTCAAGTGAAGAGAAACCTGTGAGATTTAACTGAGAT 1936
 Db 1658 ATGAGAGCCCAAGAAATATCTCAAGTGAAGAGAAACCTGTGAGATTTAACTGAGAT 1717
 Qy 1937 ATGAGAGATCCCGCAAAATCACTCAAGAAAGCTTTGAGAACTGCAAGCAAGCTTA 1996
 Db 1718 ATGAGAGATCCCGCAAAATCACTCAAGAAAGCTTTGAGAACTGCAAGCAAGCTTA 1777
 Qy 1997 CACACATCACTCAAGATCAATGATCTTATTTGGGCTTGAAGAAACAGGATGAGCA 2056
 Db 1778 CACACATCACTCAAGATCAATGATCTTATTTGGGCTTGAAGAAACAGGATGAGCA 1837
 Qy 2057 CATTCCTAACTGCTTCTCAATGAGGAGTGTGAGGAGGCTTCACTGCTGATG 2116
 Db 1838 CATTCCTAACTGCTTCTCAATGAGGAGTGTGAGGAGGCTTCACTGCTGATG 1897
 Qy 2117 ACCAGAGAGCACTGATCAAGTCAAGCAATGAGTGTGCACTGAGAGAGAGTGTCT 2176
 Db 1898 ACCAGAGAGCACTGATCAAGTCAAGCAATGAGTGTGCACTGAGAGAGAGTGTCT 1957
 Qy 2177 ACAAGAGCTGAATCAAAATGCTTGAAGTGAATGAATGAATGAATGAATGAATGAAT 2236
 Db 1958 ACAAGAGCTGAATCAAAATGCTTGAAGTGAATGAATGAATGAATGAATGAATGAAT 2017
 Qy 2237 TAAAGCTTTTATCATTAATAAAGTGAAGCTGAGTGAAGCTGAGTGAAGCTTAACT 2296
 Db 2018 TAAAGCTTTTATCATTAATAAAGTGAAGCTGAGTGAAGCTGAGTGAAGCTTAACT 2077
 Qy 2297 CAGACTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2356
 Db 2078 CAGACTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2137
 Qy 2357 TTAAGTATGCTGCAAGCTAATGAGTTTAATAA 2390
 Db 2138 TTAAGTATGCTGCAAGCTAATGAGTTTAATAA 2171

RESULT 2
 US-09-949-016-16644/c
 ; Sequence 16644, Application us/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.

```

? TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
? TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
? FILE REFERENCE: CLO01307
? CURRENT APPLICATION NUMBER: US/09/949,016
? CURRENT FILING DATE: 2000-04-14
? PRIOR APPLICATION NUMBER: 60/241,755
? PRIOR FILING DATE: 2000-10-20
? PRIOR APPLICATION NUMBER: 60/237,768
? PRIOR FILING DATE: 2000-10-03
? PRIOR APPLICATION NUMBER: 60/231,498
? PRIOR FILING DATE: 2000-09-08
? NUMBER OF SEQ ID NOS: 207012
? SOFTWARE: FASTSEQ FOR WINDOWS VERSION 4.0
? SEQ ID NO 16644
? LENGTH: 7609
? TYPE: DNA
? ORGANISM: Human
? OS-09-949-016-16644

```

Query March	26.4%	Score 642.2	DB 4	Length 7609
Beef Local	99.5%	Pred. No. 4e-157		
Matches 644	Conservative 0	Mismatches 3	Indels 0	Gaps 0

Qy	1405	CAGCCTGAACTGAGGAGGAGTTCACGCCCAATCAGAGAGCAAGAGATGCTCA	1464
Db	7382	CAGCTGAACAAAGAGAGATGAGATTTCACGCCCAATCAGAGAGCAAGAGATGCTCA	7323
Qy	1465	CGCTGGCATATTGCGCGCTTCTTAGGCGCTGCACTCCATATCAACATCTCCAGGGC	1524
Db	7322	CGCTGGCATATTGCGCGCTTCTTAGGCGCTGCACTCCATATCAACATCTCCAGGGC	7263
Qy	1525	AAACTCAAGGACCAAACTACTGTGGGCCCAAGAGGAAATCCCTTCCACAGAGGCGTGGC	1586
Db	7262	AAACTCAAGGACCAAACTACTGTGGGCCCAAGAGGAAATCCCTTCCACAGAGGCGTGGC	7203
Qy	1585	CAAAAACCAACAGGCGAGCCAAACAGAACGTTTAGGGGCCAGAGAGCAACAAGGCGCTGGAA	1644
Db	7202	CAAAAACCAACAGGCGAGCCAAACAGAACGTTTAGGGGCCAGAGAGCAACAAGGCGCTGGAA	7143
Qy	1645	GCTTAAGGCTGTGAGCGCCCTTCAAGTCTGGGCCCACTGTATAGAGGCGTAGGCTACTACAG	1704
Db	7142	GCTTAAGGCTGTGAGCGCCCTTCAAGTCTGGGCCCACTGTATAGAGGCGCTACTACAG	7083
Qy	1705	TGCCCCACAGCGGCCCTCAGGCCCACTGCCATGTTTCCGCCTAGAACCATACAGGCCC	1766
Db	7082	TGCCCCACAGCGGCCCTCAGGCCCACTGCCATGTTTCCGCCCTAGAACCATACAGGCCC	7023
Qy	1765	GTCAAAGCTTACACAGTGTCAAGGCATAGACCAACAAAGACTTAAGAAGCTGTGAG	1824
Db	7022	GTCAAAGCTTACACAGTGTCAAGGCATAGACCAACAAAGACTTAAGAAGCTGTGAG	6965
Qy	1825	TTCTGGGGCCAAAGAAAGTTTGAATTGCTCTCAGAGAGCGATGGGGCTTGATGGAGCA	1884
Db	6962	TTCTGGGGCCAAAGAAAGTTTGAATTGCTCTCAGAGAGCGATGGGGCTTGATGGAGCA	6903
Qy	1885	CCCAAGAGTATCTCAAGTGAAGAGGAAACTGTGAGTTTAACCTGAACGATATGCGCAGA	1944
Db	6902	CCCAAGAGTATCTCAAGTGAAGAGGAAACTGTGAGTTTAACCTGAACGATATGCGCAGA	6843
Qy	1945	GATCCCGGAAATCACTCAAGAAACCTTTGGAACAATCAACCAACCAATACACATACACTAC	2004
Db	6842	GATCCCGGAAATCACTCAAGAAACCTTTGGAACAATCAACCAACCAATACACATACACTAC	6783
Qy	2005	ACTCAAGATCACTGATCTCTTATTTGGGCTTTGGAAACAGGGATC 2051	
Db	6782	ACTCAAGATCACTGATCTCTTATTTGGGCTTTGGAAACAGGGATC 6736	

```

RESULT 3
US-09-949-016-16643
; Sequence 16643, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:

```

```

APPLICANT : ENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH HUMAN DISEASE. METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT FILING DATE: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 16643
LENGTH: 18075
TYPE: DNA
ORGANISM: Human
US-09-949-016-16643

```

Query Match	26.4%	Score 642.2;	DB 4;	Length 18075;
Best Local Similarity	99.5%;	Pred. No. 6.6e-157;		
Matches 644;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0

OY	1405	CAGCTGAA	CAAA	GAGAGATG	AGTTCC	AGCCCAAT	CAGAGAGAA	CAAGAGATG	CTCA	1464	
Db	13879	CAGCTGAA	CAAA	GAGAGATG	AGTTCC	AGCCCAAT	CAGAGAGAA	CAAGAGATG	CTCA	13938	
OY	1465	CGCTGGAT	CATATT	GCGCGCTT	CTTAGG	CGTCGAGT	CCATGAT	CAACAT	CTCTCC	1524	
Db	13939	CGCTGGAT	CATATT	GCGCGCTT	CTTAGG	CGTCGAGT	CCATGAT	CAACAT	CTCTCC	13988	
OY	1525	AAACTCA	AGGAC	CCAACTA	CTGTGG	CCCAAGGGAA	TTCCTT	CTCCAC	GAGGCTG	1584	
Db	13999	AAACTCA	AGGAC	CCAACTA	CTGTGG	CCCAAGGGAA	TTCCTT	CTCCAC	GAGGCTG	14058	
OY	1585	CAAAAA	AC	CAAGAG	CGCCAA	CAGAA	CGTTAG	GGGCGC	AGAAAG	CAACAG	1644
Db	14059	CAAAAA	AC	CAAGAG	CGCCAA	CAGAA	CGTTAG	GGGCGC	AGAAAG	CAACAG	14118
OY	1645	GCTTTA	AGGCTGT	GAGCG	CGCTT	CAAGTGT	GAGCCCA	CTGTAT	CAGAGCG	CGCTACT	1704
Db	14119	GCTTTA	AGGCTGT	GAGCG	CGCTT	CAAGTGT	GAGCCCA	CTGTAT	CAGAGCG	CGCTACT	14178
OY	1705	TGCCCCA	CAGA	CGCCCT	CAGCCCA	CTCCAT	GTCTT	CCCCCT	TAGAAC	CAATC	1764
Db	14179	TGCCCCA	CAGA	CGCCCT	CAGCCCA	CTCCAT	GTCTT	CCCCCT	TAGAAC	CAATC	14238
OY	1765	GTCAAAG	CTTCA	CAGTGT	CA	CAGGCA	CTTAA	CAAC	CAAAAG	CTTAA	1824
Db	14239	GTCAAAG	CTTCA	CAGTGT	CA	CAGGCA	CTTAA	CAAC	CAAAAG	CTTAA	14298
OY	1825	TTCGTGG	GGCC	CAAGAAA	GTTTG	AAATTG	CTCT	CA	GAGAG	CGATG	1884
Db	14299	TTCGTGG	GGCC	CAAGAAA	GTTTG	AAATTG	CTCT	CA	GAGAG	CGATG	14358
OY	1885	CCGAGA	AGAT	CTCA	AGTGA	GAGGAAA	CTGTG	AGGTTAA	CTGAG	GGATG	1944
Db	14359	CCGAGA	AGAT	CTCA	AGTGA	GAGGAAA	CTGTG	AGGTTAA	CTGAG	GGATG	14418
OY	1945	GATCCCC	GAAAT	CACCT	CA	AGAA	GCCTT	TGAA	CAAT	CAAC	2004
Db	14419	GATCCCC	GAAAT	CACCT	CA	AGAA	GCCTT	TGAA	CAAT	CAAC	14478
OY	2005	ACTCAAA	GATCA	CTGAT	CTTATT	TGGG	CTTTG	GA	AAAC	GAGGATG	2051
Db	14479	ACTCAAA	GATCA	CTGAT	CTTATT	TGGG	CTTTG	GA	AAAC	GAGGATG	14525

RESULT 4
US-09-949-016-4902/c
; Sequence 4902, Application US/09949016
; Patent No. 6812339

```
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO: 4902
/ LENGTH: 1462
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-4902
```

```
Query Match      14.8%; Score 358.4; DB 4; Length 1462;
Best Local Similarity 99.7%; Pred. No. 2.2e-83;
Matches 359; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 2045 AGGATGAGCAGCATTTCTTAACCTGCTTCTTAATGGGAGATGCTTGGCAGCCAGGTCCTC 2104
DB 1385 AGGATGAGCAGCATTTCTTAACCTGCTTCTTAATGGGAGATGCTTGGCAGCCAGGTCCTC 1326
QY 2105 ACTGTGTGTACACAGAGAGACATGATCCAGTCACAGCCATACAGCTGTCCACACTGA 2164
DB 1325 ACTGTGTGTACACAGAGAGACATGATCCAGTCACAGCCATACAGCTGTCCACACTGA 1266
QY 2165 AGAAGCTGTCTTCAACAAGCCTGAATCAATGTTAGTTAATAGATAAAATCCAGAC 2224
DB 1265 AGAAGCTGTCTTCAACAAGCCTGAATCAATGTTAGTTAATAGATAAAATCCAGAC 1206
QY 2225 TACTTCAGCCTTTAATGCTTTTATTCATTAATAAAGCTGTAAGCTGAACCATTTG 2284
DB 1205 TACTTCAGCCTTTAATGCTTTTATTCATTAATAAAGCTGTAAGCTGAACCATTTG 1146
QY 2285 AAACATTTAACCTGAGCTGTGATTCAGAGTGGGAAACCTTAGTTCTATCGAATCCAA 2344
DB 1145 AAACATTTAACCTGAGCTGTGATTCAGAGTGGGAAACCTTAGTTCTATCGAATCCAA 1086
QY 2345 GACAGCCACACCTTAGTATCTGCCCAACTAATGATTAAATAACAATACTCGTT 2404
DB 1085 GACAGCCACACCTTAGTATCTGCCCAACTAATGATTAAATAACAATACTCGTT 1026
```

```
RESULT 5
US-09-949-016-173056
/ Sequence 173056, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO: 173056
/ LENGTH: 601
/ TYPE: DNA
/ ORGANISM: Human
```

```
US-09-949-016-173056
```

```
Query Match      7.8%; Score 188.4; DB 4; Length 601;
Best Local Similarity 99.5%; Pred. No. 3.3e-39;
Matches 189; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 2045 AGGATGAGCAGCATTTCTTAACCTGCTTCTTAATGGGAGATGCTTGGCAGCCAGGTCCTC 2104
DB 412 AGGATGAGCAGCATTTCTTAACCTGCTTCTTAATGGGAGATGCTTGGCAGCCAGGTCCTC 471
QY 2105 ACTGTGTGTACACAGAGAGACATGATCCAGTCACAGCCATACAGCTGTCCACACTGA 2164
DB 472 ACTGTGTGTACACAGAGAGACATGATCCAGTCACAGCCATACAGCTGTCCACACTGA 531
QY 2165 AGAAGCTGTCTTCAACAAGCCTGAATCAATGTTAGTTAATAGATAAAATCCAGAC 2224
DB 532 AGAAGCTGTCTTCAACAAGCCTGAATCAATGTTAGTTAATAGATAAAATCCAGAC 591
QY 2225 TACTTCAGCC 2234
DB 592 TACTTCAGCC 601
```

```
RESULT 6
US-09-949-016-173051
/ Sequence 173051, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO: 173051
/ LENGTH: 601
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-173051
```

```
Query Match      6.6%; Score 161.4; DB 4; Length 601;
Best Local Similarity 99.4%; Pred. No. 3.8e-32;
Matches 162; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 582 GCAGGCTTTATGACTCCGAGAACACAGAGTTGAGAACTGAGCTTACACACAC 641
DB 321 GCAGGCTTTATGACTCCGAGAACACAGAGTTGAGAACTGAGCTTACACACAC 380
QY 642 ATGTTCTGAGTCCCTGGGAGTGTGTTGCAACCTGTCAATGAAGCGTGTGAGGT 701
DB 381 ATGTTCTGAGTCCCTGGGAGTGTGTTGCAACCTGTCAATGAAGCGTGTGAGGT 440
QY 702 CGAATCCGGAGCCCTATCTGTCTCCAGAGCCTGTGAACGAGA 744
DB 441 CGAATCCGGAGCCCTATCTGTCTCCAGAGCCTGTGAACGAGA 483
```

```
RESULT 7
US-09-949-016-173052
/ Sequence 173052, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
```


Qy 1141 GAGAATCGCTTGAAACCCGGAGCGAGGTTGTGGTGCGAGAGCAGCTCATCA 1193
|||||
Db 17739 TAGAATCGCTTGAAACCCGGAGCGAGGTTGTGGTGAGCCCAAGATCACACCA 17687

RESULT 11
US-09-949-016-46055/C
; Sequence 46055, Application US/09949016

```

: GENERAL INFORMATION:
: APPLICANT: VENTER, J. Craig et al.
: TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
: TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: C1001107
: CURRENT APPLICATION NUMBER: US/09/949,016
: CURRENT FILING DATE: 2000-04-14
: PRIOR APPLICATION NUMBER: 60/241,755
: PRIOR FILING DATE: 2000-10-20
: PRIOR APPLICATION NUMBER: 60/237,768
: PRIOR FILING DATE: 2000-10-03
: PRIOR APPLICATION NUMBER: 60/231,498
: PRIOR FILING DATE: 2000-09-08
: NUMBER OF SEQ ID NOS: 207012
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 46055
:
: LENGTH: 601
:
: TYPE: DNA
: ORGANISM: Human
: US-09-949-016-46055

```

[illegible]

```

RESULT 12
US-09-949-016-13059/C
/ Sequence 13059, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CLO01307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: fastseq for Windows Version 4.0
/ SEQ ID NO 13059
/ LENGTH: 43086
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-13059

```

Query Match	3.6%	Score 88.6;	DB 4;	Length 43086;
Best Local Similarity	91.3%;	Pred. No. 4.5e-12;		
Matches 94;	Conservative 0;	Mismatches 9;	Indels 0;	Gaps 0;

QY 1081 AAACACACAGGGGTGGTGGCGACACCTGTATCCCACTCTCTGGAGGCTGAGGCAG 1140

Db 10245 AATATTACCGGGCGGTGTGGCGATCTGTATCCCACTATGTGGGGGGCTGAGGCAG 10186

QY 1141 GAGCAATGCTTTGAACCGGGAGCGGAGGTTGTGTGGCGAGG 1183

Db 10185 GAGATATGCTTTGAACCGGGAGCGGAGGTTGTGTGGCGAGG 10143

RESULT 13
US-09-949-016-14577

```

; Patent No.6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14577
; LENGTH: 678533
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(678533)
; OTHER INFORMATION: n = A,T,C or G
; US-09-949-016-14577

```

	Query Match	3.6%	Score 88.4	DB 4	Length 678553
	Best Local Similarity	73.4%	Pred. No. 2.5e-11		
	Matches 113	Conservative	0	Mismatches 41	Indels 0
				Gaps	0
Qy	1079	AGAAAGCAGCAGGGGTGTGGCGACACTGTATCTCCAGCTACTCGGAGGCTGAGCC	1138		
Db	499918	AAAAAAGCTGCGAGCGGTGTGGCGACCCCTGTATCTCCAGCTACTCAGAGTGTGAGAC	499977		
Qy	1139	AGGAGAAATCGCTTGAACCCGCGAGGTTGTGTGGCGACGACGCTCATCAACCC	1198		
Db	499978	AGGAGAAATCGCTTGAACCCGCGAGGTTGTGTGGCGACGACGATCATCAACGCC	500037		
Qy	1199	TTTGGAGAGAGTATGATGATTTTGAACCAACT	1232		
Db	500038	ATTGACTCCAGCTTGGGCAACAGAGCAAACT	500071		

RESULT 14
US-09-949-016-14578
: Sequence 14578, Application US/09949016
: Patent No. 6012339
: GENERAL INFORMATION:
: APPLICANT: VENTER, J. Craig et al.
: TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
: TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: C0001307
: CURRENT APPLICATION NUMBER: US/09/949,016
: CURRENT FILING DATE: 2000-04-14
: PRIOR APPLICATION NUMBER: 60/241,755
: PRIOR FILING DATE: 2000-10-20
: PRIOR APPLICATION NUMBER: 60/237,768
: PRIOR FILING DATE: 2000-10-03
: PRIOR APPLICATION NUMBER: 60/231,498
: PRIOR FILING DATE: 2000-09-08
: NUMBER OF SEQ ID NOS: 207012

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2005, 22:00:09 ; Search time 8160.63 Seconds
(without alignments)
11376.592 Million cell updates/sec

Title: US-09-622-964A-28
Perfect score: 1916
Sequence: 1 gtgcgaagcatgcatcata.....aaaaaaaaaaaaaaaa 1916

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1916	100.0	1916	6	BD136744 Best's ma
2	1874.8	97.8	1904	10	AY450427 Mus muscu
3	1331.4	69.5	2035	10	BC079048 Rattus no
4	873.8	45.6	2187	9	AY357925 Macaca fa
5	872.8	45.6	2171	9	AF073501 Homo sapi
6	871.2	45.5	2210	6	AF057169 Homo sapi
7	871.2	45.5	2229	6	BD136720 Best's ma
8	867	45.3	1758	6	AX745964 Sequence
9	865.4	45.2	1758	9	AY15704 Homo sapi
10	859.2	44.8	2170	6	CQ729633 Sequence
11	777.2	40.6	2435	9	BC041664 Homo sapi
12	658.2	34.4	2420	9	AF057170 Homo sapi
13	658.2	34.4	2429	6	BD136721 Best's ma
14	624.4	32.6	2441	9	BC015820 Homo sapi
15	519.4	27.1	2137	6	AX745966 Sequence
16	519	27.1	1530	6	CQ731444 Sequence
17	519	27.1	1530	6	AX745993 Sequence
18	519	27.1	1530	9	AY15705 Homo sapi
19	519	27.1	1908	9	AF440756 Homo sapi

20	491.2	25.6	1289	4	AY064707 Sus scrofa
21	486.8	25.4	1861	9	AK000139 Homo sapi
22	479.4	25.0	2719	10	AY450426 Mus muscu
23	479	25.0	1954	10	AY450428 Mus muscu
24	479	25.0	1956	10	BC031186 Mus muscu
25	479	25.0	1957	10	BC019528 Mus muscu
26	479	25.0	1992	10	BC036157 Mus muscu
27	479	25.0	2028	10	BC036163 Mus muscu
28	458.4	23.9	1506	9	AF440758 Homo sapi
29	458.2	23.9	2007	9	AY15706 Homo sapi
30	458.2	23.9	2500	6	AX834175 Sequence
31	458.2	23.9	2500	9	AK096459 Homo sapi
32	445.8	23.3	1898	5	BC043854 Xenopus t
33	445.8	23.3	1921	5	AY273825 Xenopus t
34	433.4	22.6	2064	5	BC084329 Xenopus t
35	431.8	22.5	1650	5	AY273826 Xenopus t
36	423	22.1	2028	6	AX745968 Sequence
37	422.2	22.0	1882	5	CR760914 Xenopus t
38	422.2	22.0	2021	5	BC061379 Sequence
39	399.2	20.8	1422	6	AX745970 Sequence
40	399.2	20.8	1422	9	AY15707 Homo sapi
41	399.2	20.8	2045	9	AF440757 Homo sapi
42	357	18.6	171749	10	AC134437 Homo sapi
43	296.2	15.5	2861	6	CQ585672 Sequence
44	296.2	15.5	2874	3	AF218817 Drosophila
45	296.2	15.5	5610	3	AY061546 Drosophila

ALIGNMENTS

RESULT 1
LOCUS BD136744 1916 bp DNA linear PAT 18-SEP-2002
DEFINITION Best's macular dystrophy gene.
ACCESSION BD136744.1 GI:23231689
VERSION JP 2002504559-A/26.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1 (bases 1 to 1916)
Petrushkin,K., Caskey,T.C., Metzker,M. and Wadelius,C.
Best's macular dystrophy gene
Patent: JP 2002504559-A 26 12-FEB-2002;
JOURNAL MERCK & CO INC, CLAS WADLIUS
COMMENT OS Mus musculus (mouse)
PN JP 2002504559-A/26
PD 12-FEB-2002
PP 22-FEB-1999 JP 200053447
PR 25-FEB-1998 US 60/075941,18-DEC-1998 US 60/112926 PI
KONSTANTIN PETRUSHKIN, THOMAS C CASKEY, MICHAEL METZKER, CLAES PI

WADLIUS
PC C07K16/18, C07K14/47, C12N5/10, C12N15/09, C12P19/34, C12Q1/68// PC
C12P21/08,
PC C12N5/00, C12N15/00
CC Best's macular dystrophy gene
FH Key location/Qualifiers
FT source 1..1916
FT location/Qualifiers
FT /organism="Mus musculus (mouse)".

FEATURES
Source location/Qualifiers
1..1916
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"

ORIGIN

Query Match 100.0%; Score 1916; DB 6; Length 1916;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1916; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 GTGCAAGCATGATCATACCTACCAACAAAGTACCAATGCCGCGCTGCTTCTT 60

Db	1	GTCCAAAGCCAGTACTATCACTACAGAAACAAAGTAGCCAAATCCCGCTTCGTTGCTT	60
Qy	61	CTCGCCCTCCCTCCGTGTCGTCGAGGAGAGATCTACAGAGTCTGTATGAGAAATTCCT	120
Db	61	CTCGCCCTCCCTCCGTGTCGTCGAGGAGAGATCTACAGAGTCTGTATGAGAAATTCCT	120
Qy	121	TGTCTTCAATATCTCTACTATTCATCCGTCGACTCAAGAAATGGTCTCTCGAGTGA	180
Db	121	TGTCTTCAATATCTCTACTATTCATCCGTCGACTCAAGAAATGGTCTCTCGAGTGA	180
Qy	181	TCACAGCTGTGTTTGAAGAGCTGTCGTACAGCAAGCTCACTTCAAGTCTTCC	240
Db	181	TCACAGCTGTGTTTGAAGAGCTGTCGTACAGCAAGCTCACTTCAAGTCTTCC	240
Qy	241	TATATCCCTGCTTGGGTTTCTATGTTACATGGTGTGAGCCGCTGTGTGAGCCAGTA	300
Db	241	TATATCCCTGCTTGGGTTTCTATGTTACATGGTGTGAGCCGCTGTGTGAGCCAGTA	300
Qy	301	CGAAGACTTCCCGTGGCCGACCGCTCATGATCCAGTGTAGCTTCTGTGAGGAGCA	360
Db	301	CGAAGACTTCCCGTGGCCGACCGCTCATGATCCAGTGTAGCTTCTGTGAGGAGCA	360
Qy	361	GGATGAGAGAGGCGTTTGTGGGGGCGAGCTCATCCGCTACGCCATCTGTGGCCAACT	420
Db	361	GGATGAGAGAGGCGTTTGTGGGGGCGAGCTCATCCGCTACGCCATCTGTGGCCAACT	420
Qy	421	GCTCATCTGCGCAGCATGAGACCTCGGTCTACAGCCCTTCCACCTTCCACACCT	480
Db	421	GCTCATCTGCGCAGCATGAGACCTCGGTCTACAGCCCTTCCACCTTCCACACCT	480
Qy	481	GGTGTAGAGAGGTTTATGACCAATGGGAAATTAAGCAGTTGCAAGATTGGGCTTACC	540
Db	481	GGTGTAGAGAGGTTTATGACCAATGGGAAATTAAGCAGTTGCAAGATTGGGCTTACC	540
Qy	541	ACAACAACATTTCTGGGTCCTTGGGTGTGTTTGCAATGTGCAATGAGAGGCTTACT	600
Db	541	ACAACAACATTTCTGGGTCCTTGGGTGTGTTTGCAATGTGCAATGAGAGGCTTACT	600
Qy	601	TGGAGTGCATATCCGGAGACCGCTTCGTCCAGAGCTGATGATGAGTGTGATCTT	660
Db	601	TGGAGTGCATATCCGGAGACCGCTTCGTCCAGAGCTGATGATGAGTGTGATCTT	660
Qy	661	GGTACTAGTGTGACAGTGTATGCTGATGACCTGATGATGATGATGATGATGATGAT	720
Db	661	GGTACTAGTGTGACAGTGTATGCTGATGACCTGATGATGATGATGATGATGATGAT	720
Qy	721	ACAGGTGTGACAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT	780
Db	721	ACAGGTGTGACAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT	780
Qy	781	TCTGAACCAACAAAGAGCTACCCAGGCAATGATGATGATGATGATGATGATGATGAT	840
Db	781	TCTGAACCAACAAAGAGCTACCCAGGCAATGATGATGATGATGATGATGATGATGAT	840
Qy	841	AATCTGCAATTTCTATCTACATGAGCTGCTGATGATGATGATGATGATGATGATGAT	900
Db	841	AATCTGCAATTTCTATCTACATGAGCTGCTGATGATGATGATGATGATGATGATGAT	900
Qy	901	CTTGGGAGAGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	960
Db	901	CTTGGGAGAGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	960
Qy	961	GTCCTGTGTCCTGTGATGAGTACAGCAAGACTTGCCTCCATGAGAACTGATGATGAT	1020
Db	961	GTCCTGTGTCCTGTGATGAGTACAGCAAGACTTGCCTCCATGAGAACTGATGATGAT	1020
Qy	1021	CTGAAAGAGAGAGAGCTTACGCTGCTTCTGCAAGGCTTCTGCGGAGCA	1080
Db	1021	CTGAAAGAGAGAGAGCTTACGCTGCTTCTGCAAGGCTTCTGCGGAGCA	1080
Qy	1081	TTCCCTCATGGGCTTCAACCTTCAACCTTCAACCTTCAACCTTCAACCTTCAACCTT	1140
Db	1081	TTCCCTCATGGGCTTCAACCTTCAACCTTCAACCTTCAACCTTCAACCTTCAACCTT	1140
Qy	1141	AAAAGAGAGGCTGACAGATTAAGAAAGAGGCTGATGAGGACCTTAGGCTCTT	1200
Db	1141	AAAAGAGAGGCTGACAGATTAAGAAAGAGGCTGATGAGGACCTTAGGCTCTT	1200
Qy	1201	CTTAGAGTCAACCCCAAACTACCATCTTCCCTTGAAGAGCTTAAAGCAACTATT	1260
Db	1201	CTTAGAGTCAACCCCAAACTACCATCTTCCCTTGAAGAGCTTAAAGCAACTATT	1260
Qy	1261	GTCCTTCAAGAACCCCTCCCTCCAGAGGCTGATGATGATGATGATGATGATGATGAT	1320
Db	1261	GTCCTTCAAGAACCCCTCCCTCCAGAGGCTGATGATGATGATGATGATGATGATGAT	1320
Qy	1321	AGATGTCTGAAATTTAAGGCTGAGCTTCTGAAATGTGTTCAAGGTTTAAAGGAG	1380
Db	1321	AGATGTCTGAAATTTAAGGCTGAGCTTCTGAAATGTGTTCAAGGTTTAAAGGAG	1380
Qy	1381	AGGCTCCCATTTGTGGCCCAAGGACCCAGAGCCCACTTACAGAGTCAAGACCTTC	1440
Db	1381	AGGCTCCCATTTGTGGCCCAAGGACCCAGAGCCCACTTACAGAGTCAAGACCTTC	1440
Qy	1441	CAGTTCAAGACAGAGTATGAGGCTTCAAGATTAACCAAGAAATGTGTCAATGAAAA	1500
Db	1441	CAGTTCAAGACAGAGTATGAGGCTTCAAGATTAACCAAGAAATGTGTCAATGAAAA	1500
Qy	1501	GAAACTGTGAGTTTAACTTGAACATTCAGAGAGCCCAAGAACTTCAACAGCG	1560
Db	1501	GAAACTGTGAGTTTAACTTGAACATTCAGAGAGCCCAAGAACTTCAACAGCG	1560
Qy	1561	CCGTTTGAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1620
Db	1561	CCGTTTGAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1620
Qy	1621	TCCCTACAGGATGAGTGTGACCAACCTGTTCTATGATGATGATGATGATGATGAT	1680
Db	1621	TCCCTACAGGATGAGTGTGACCAACCTGTTCTATGATGATGATGATGATGATGAT	1680
Qy	1681	GCCCTGATTTGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1740
Db	1681	GCCCTGATTTGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1740
Qy	1741	ACACCCAGAGAGTGTGTTCCAGAGAGTGTGATGATGATGATGATGATGATGATGAT	1800
Db	1741	ACACCCAGAGAGTGTGTTCCAGAGAGTGTGATGATGATGATGATGATGATGATGAT	1800
Qy	1801	TAGTCTGCTGTAACCAACCTGATTTTGAAGATCTTCCCAACTTAAGATTTAATAA	1860
Db	1801	TAGTCTGCTGTAACCAACCTGATTTTGAAGATCTTCCCAACTTAAGATTTAATAA	1860
Qy	1861	CCTGAATATTTCTTTAGGTGAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAA	1916
Db	1861	CCTGAATATTTCTTTAGGTGAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAA	1916
RESULT 2			
AY450427 1904 bp mRNA linear ROD 24-JUN-2004			
LOCUS AY450427			
DEFINITION Mus musculus vitelliniform macular dystrophy 2 (Vmd2) mRNA, complete			
ACCESSION AY450427			
VERSION AY450427.1 GI:41612988			
KEYWORDS Mus musculus (house mouse)			
SOURCE ORGANISM			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
REFERENCE 1 (bases 1 to 1904)			
AUTHORS Stohr, H., Marquardt, A., Nanda, I., Schmid, M. and Weber, B. H.			
TITLE Three novel human VMD2-like genes are members of the evolutionary			
JOURNAL Eur. J. Hum. Genet. 10 (4), 261-264 (2002)			
MEDLINE 22027749			

PUBMED 12032738
REFERENCE 2 (bases 1 to 1904)
AUTHORS Kraemer, F., Stoeck, H., and Weber, B.H.F.
TITLE Cloning and characterization of the murine Vmd2 Ref-1m gene family
JOURNAL Cytogenet. Genome Res. 105 (1), 107-114 (2004)
RECEIVED 3 (bases 1 to 1904)
AUTHORS Weber, B.H.F., and Kraemer, F.
TITLE Direct Submission
JOURNAL Submitted (27-OCT-2003) Human Genetics, University of Wuerzburg, Am
Hubland, Biozentrum, Wuerzburg 97074, Germany
FEATURES
Source Location/Qualifiers
1..1904
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/chromosome="19"
/tissue_type="testis"
1..1904
/gene="Vmd2"
37..1692
/gene="Vmd2"
/codon_start=1
/product="Vitellogenin macular dystrophy 2"
/protein_id="AA05922.1"
/db_xref="GI:41612989"
/translation="MTITVTMKVANARLGSFSSLLCMRGSITKYLIVGEPLVFLVY
SIRLRYMVLSSDQQLPEKALYCDSTIQIPISPVIGFVTVTVVSRMMSQYENLPM
PRLMIVQSFVEKDEBGRLLRLIRLALQVLLKRSISVYKRPFLHLVLA
GPMTHGKQKSLGKLPNTFWVWVPMFANLSMAYGRLRDVLLQSLMNEVCTLR
TQCCQAYADWISILVYTVVAVVVSFLACILGKQFLPNKYDYGHEMLVVPF
TILQFLPMGLKVAEOLINFGEDDDFEFTNMLIDNRLQVSLSVQGHONLPMER
DMVNEAAROPPYTAARSARSRHSFMSTFISLKEDLDELSEKADTLKKESGYSS
TIGFELQIPKNTYLPKDLTKILCSNPLLEBQCCDANKXNDKPKRGLDPLK
VPRKRSHSCGPAPSPHTEGSPSSSDTGDPSTDYQSGIMKKKTVEFNLTIPS
SPTEHLQORLDQSTNLIQALMKEHASYPRDEAGTKPVLYE"

ORIGIN
Query Match 97.8%; Score 1874.8; DB 10; Length 1904;
Base Local Similarity 99.9%; Pred. No. 0;
Matches 1876; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Query 1 GTGCCAAGCCATGATCATCTACACAAACAAAGTACCAATGCCCGCTCGGTTCTT 60
Db 27 GTGCCAAGCCATGATCATCTACACAAACAAAGTACCAATGCCCGCTCGGTTCTT 86
Query 61 CTGCTCCCTCTCTGTGCTGCGAGGAGCATCTACAAGCTGCTGTAATGGAATTCCT 120
Db 87 CTGCTCCCTCTCTGTGCTGCGAGGAGCATCTACAAGCTGCTGTAATGGAATTCCT 146
Query 121 TGTCTTCAATTTCTCTACTATTCATTCCTGGGACTCTACAGAAATGTTCTTCTGAGTGA 180
Db 147 TGTCTTCAATTTCTCTACTATTCATTCCTGGGACTCTACAGAAATGTTCTTCTGAGTGA 206
Query 181 TCACAGAGCTGTTGTTGGAAGCTGAGCTCTGTAACAGACAGTACATTCAGCTCATCC 240
Db 207 TCACAGAGCTGTTGTTGGAAGCTGAGCTCTGTAACAGACAGTACATTCAGCTCATCC 266
Query 241 TATAATCTTCTGTTCTGAGTTTCTATGTTACATTTGGTGTGAAGCCGCTGGTGAAGCAGTA 300
Db 267 TATAATCTTCTGTTCTGAGTTTCTATGTTACATTTGGTGTGAAGCCGCTGGTGAAGCAGTA 326
Query 301 CGAATACTTGGCGGTGGCCGACCGCTCATATATCCAGGTGTCTAGCTTCTGGAAGGGGA 360
Db 327 CGAATACTTGGCGGTGGCCGACCGCTCATATATCCAGGTGTCTGGAAGGGGA 386
Query 361 GGATGAGGAAGAGCGTTTGTGCGGCGGACAGCTCATCCGCTACAGCATTCCTGGGCGAAGT 420
Db 387 GGATGAGGAAGAGCGTTTGTGCGGCGGACAGCTCATCCGCTACAGCATTCCTGGGCGAAGT 446
Query 421 GCTCATCTGCGAGCATCAGCACTGGGTCTACAAAGCGCTTTCCCATCTTCAACCACT 480
Db 447 GCTCATCTGCGAGCATCAGCACTGGGTCTACAAAGCGCTTTCCCATCTTCAACCACT 506

Query 481 GGTGCTAGCAGTTTATGACCCATGGGGAACATTAACAGTTTGAGAAATTGGGCTTACC 540
Db 507 GGTGCTAGCAGTTTATGACCCATGGGGAACATTAACAGTTTGAGAAATTGGGCTTACC 566
Query 541 ACACAACATTTCTGGGTGCCCTGGGTGTGTTGGCCATTTGCAATGAAGGCTATCT 600
Db 567 ACACAACATTTCTGGGTGCCCTGGGTGTGTTGGCCATTTGCAATGAAGGCTATCT 626
Query 601 TGGAGTGGAAATCGGGGACACCGCTGCTCAGAGCTGATGAATGAGGTGTACTT 660
Db 627 TGGAGTGGAAATCGGGGACACCGCTGCTCAGAGCTGATGAATGAGGTGTACTT 686
Query 661 GCGTACTCAGTGTGAGACAGCTGTATGCTTACGACTGTATTAATGCCATTTGGTACAC 720
Db 687 GCGTACTCAGTGTGAGACAGCTGTATGCTTACGACTGTATTAATGCCATTTGGTACAC 746
Query 721 ACAGTGTGTACAGTGTGACATTAACAGCTTTTCTTGCATGCTTGAATCGGAGGAGTT 780
Db 747 ACAGTGTGTACAGTGTGACATTAACAGCTTTTCTTGCATGCTTGAATCGGAGGAGTT 806
Query 781 TCTGAACCCAAACAAAGACATACCCAGGCGCATGAGATGAGTCTGCTGTCTTAC 840
Db 807 TCTGAACCCAAACAAAGACATACCCAGGCGCATGAGATGAGTCTGCTGTCTTAC 866
Query 841 AATCCTCAATTTCTTATTTCTACATGAGCTGGCTGAAGGTGACAGACAGCTCATCAACC 900
Db 867 AATCCTCAATTTCTTATTTCTACATGAGCTGGCTGAAGGTGACAGACAGCTCATCAACC 926
Query 901 CTTCGGGAGAGACGATGATGATTTTGAAGTAACTGATCATTAACAGAAACTTGCAGGT 960
Db 927 CTTCGGGAGAGACGATGATGATTTTGAAGTAACTGATCATTAACAGAAACTTGCAGGT 986
Query 961 GTCCCTGTGCTGCTGATGATGATGACCAAGAACTTGCCTCCCATGGAAGCTGACATGTA 1020
Db 987 GTCCCTGTGCTGCTGATGATGATGACCAAGAACTTGCCTCCCATGGAAGCTGACATGTA 1046
Query 1021 CTGGAACGAGCAGCGCTCAGCGGCTGACACAGCTGCTTCTGCAAGTCTCCGCGCA 1080
Db 1047 CTGGAACGAGCAGCGCTCAGCGGCTGACACAGCTGCTTCTGCAAGTCTCCGCGCA 1106
Query 1081 TTCCTTATGAGGCTTCAACCTTCAACATCAAGCTTAAAGAAAGACCTTGAAGCTTGTGTC 1140
Db 1107 TTCCTTATGAGGCTTCAACCTTCAACATCAAGCTTAAAGAAAGACCTTGAAGCTTGTGTC 1166
Query 1141 AAAAGAGAGGCTGACACGATTAAGAAAGAGAGTGTATAGACACCAATAGGCTGTT 1200
Db 1167 AAAAGAGAGGCTGACACGATTAAGAAAGAGAGTGTATAGACACCAATAGGCTGTT 1226
Query 1201 CTTAGGACTGCAACCCAAACCTACATCTTCCCTGAAAGACTTAAAGACCAATCTATT 1260
Db 1227 CTTAGGACTGCAACCCAAACCTACATCTTCCCTGAAAGACTTAAAGACCAATCTATT 1286
Query 1261 GTGTTCTTAAGAACCCCTCTCTCGAAGGCAAGTGTAAAGTGCACCAAGAAAAACAGAA 1320
Db 1287 GTGTTCTTAAGAACCCCTCTCTCGAAGGCAAGTGTAAAGTGCACCAAGAAAAACAGAA 1346
Query 1321 AGATGTCTGAAATTTAAGGCTGTGACCTTGTGAATGTGTTCGAAGGTTAAGAGAG 1380
Db 1347 AGATGTCTGAAATTTAAGGCTGTGACCTTGTGAATGTGTTCGAAGGTTAAGAGAG 1406
Query 1381 AGGCTCCCATTTGGGCCCAAGGACCCAGGACCCAGGACCCCTTCTAGAGAGTCAAGCCTTC 1440
Db 1407 AGGCTCCCATTTGGGCCCAAGGACCCAGGACCCAGGACCCCTTCTAGAGAGTCAAGCCTTC 1466
Query 1441 CAGTTTCAAGACAGAGTATGAGGCTTCTCAAGATTACCAAGAAATCTGTCAATGAATA 1500
Db 1467 CAGTTTCAAGACAGAGTATGAGGCTTCTCAAGATTACCAAGAAATCTGTCAATGAATA 1526
Query 1501 GAAATCTGTGAGTTTAATTTGAACATTCAGAGAGGCCCAAGAACATCTTCAACAGCG 1560
Db 1527 GAAATCTGTGAGTTTAATTTGAACATTCAGAGAGGCCCAAGAACATCTTCAACAGCG 1586

QY	1561	CCGTTTGAGACAGAGTGTCAACCAATATATACAGGCTCTAATGAAGAGACATGACAGTCCMA	1621
Db	1587	CCGTTTGAGACAGAGTGTCAACCAATATATACAGGCTCTAATGAAGAGAGCATGAGACTCTTA	1646
QY	1621	TCCCTTACAGAGGATGGAAGCTGGACCAACCTGTCTCTATAGTGTATGCTTCAACGCTG	1680
Db	1647	TCCCTTACAGAGGATGGAAGCTGGACCAACCTGTCTCTATAGTGTATGCTTCAACGCTG	1706
QY	1681	GCCCTGACCTTGAAGAAGATGCCAGACAGGGCACTGACCCAGTCAAAAGGCACAACAGCAGCG	1740
Db	1707	GCCCTGACCTTGAAGAAGATGCCAGACAGGGCACTGACCCAGTCAAAAGGCACAACAGCAGCG	1766
QY	1741	ACACCCAGAGAGTGTGTTCCCAAGACAGTCTAGCATGTAACTCAGAACCAAGATCTTAA	1800
Db	1767	ACACCCAGAGAGTGTGTTCCCAAGACAGTCTAGCATGTAACTCAGAACCAAGATCTTAA	1826
QY	1801	TAGTCTGCTCCGTAAGAAACACTGTATTTTATACATCTTTTCCCAACTAAGAGCTTTAATAAA	1860
Db	1827	TAGTCTGCTCCGTAAGAAACACTGTATTTTATACATCTTTTCCCAACTAAGAGCTTTAATAAA	1886
QY	1861	CGTGAATATCTTTTAAAG	1878
Db	1887	CGTGAATATCTTTTAAAG	1904

RESULT 3	BC079048	BC079048	2035 bp	mRNA	linear	ROD 03-AUG-2004
LOCUS	BC079048					
DEFINITION	Rattus norvegicus cDNA clone	MGC:93985	IMMGE:7115255	complete cds		
ACCESSION	BC079048					
VERSION	BC079048.1	GI:50927692				
KEYWORDS	MGC.					
SOURCE	Rattus norvegicus					
ORGANISM	Rattus norvegicus (Norway rat)					

REFERENCE
1 (bases 1 to 2035)
Straussberg, R. L., Feinold, F. A., Crowder, R. H., Dunn, J. S.

TITLE
 JOURNAL
 PMID
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REMARK
 COMMENT
 CONTRACT

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932
 2 (bases 1 to 2035)
 Director MGC Project.
 Direct Submission
 Submitted (02-AUG-2004) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contract: MGC brain deck

REMARK
COMMENT
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Howard Jacobs
cDNA Library Preparation: Express Genomics
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (fint)

DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcdexax11.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/INL at: <http://image.inl.gov>
Series: IRAX Plate: 183 Row: a Column: 22
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, similarity but not identity to protein.
Location/Qualifiers
1. 2035
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="MGC:93985 IMAGE:7115295"
/tissue_type="Testis, rat (Brown Norway)"
/clone_id="NIH_MGC_237"
/lab_host="DH10B"
/note="Vector: pExpress1"
153. 1805

```

/codon_start=1
/product="Unknown (protein for MG:93985)"
/protein_id="AAH79048.1"
/db_xref="GI:50927693"
/translation="MTITTYEVANRRASBECCLLRNGSTLYKLVEEPLVTFLLY
CIGURVMTLSHQOQDFEKALVYCNSTYQPIPISEFVGLFYVTLVSSMNQYENLFW
PDLIMQVSSFVFGKQKQGRLLRLIRLAIIGQVLRIRLSTVSKRFPSLHMOA
GEMTPESYKQRLRLGLPHNTFWVPMWFMALSMKRYLGGRLIRDTLTLQSLNEXCT
TQCGHLYAYDWINIPLVTVQVTVAVSFALCIIGRQFLAPLKKYRPEHEDLVYPV
TLIQPLFWMGWLKVAEQLNPFGEDDDDPEFTWITIDRLQVLSLVGGMHDPFLER
DIYANNAAPQPPYTAASGSHRQSFMSGSTFNNSLKRRQSEESSTYDKEENDVYSTG
RFGIGDSNDYHPPEKTDLLKTKLMSKDFLEQDSKNDKQKQKNDWMPKQLPAITPK
KFSFARQAPQPLGSLPMVFPQQQASASVSDTGGGSSQVQYQEAChMkKkTYEPLN
NIPECTHEQPHLDQMHTLKEPESGTFVDEKADTNVSK"
```

Query	Local Similarity	69.5%;	Score 1331.4;	DB 10;	Length 2035;
Best Local Similarity	85.3%;	Pred. No. 0;			
Matches 1677;	Conservative	0;	Mismatches 221;	Indels 63;	Gaps 12;
Qy	1	GTGCCAAGCCATGACTATACCTTACCAACAAAGTAGCCATATGCCGCTCGGTTGGTT	60		
Db	143	GTGCCAAGCCATGACTATACCTTACCAACAAAGTAGCCATATGCCGCTTGGTTCTT	202		
Qy	61	CTGCGCCCTCCTCCTGCTGCTGGCGAGGCGACATCTAACAGCTGCTGATAGGAATTCCT	120		
Db	203	CTGCGCCCTCCTCCTGCGCTGGCGAGGCGACATCTAACAGCTGCTGATAGGAATTCCT	262		
Qy	121	TGTCCTCATATTCCTCTACTATTTCCATCCGATGGACTTACAGAAATGGTCTCTCGATGA	180		
Db	263	TGTCCTCATATTCCTCTACTATTTCCATCCGATGGACTTACAGAAATGGTCTCTCGATGA	322		
Qy	181	TGAGCAAGCTGTTTGTGAGAACGTCCTGTATCTGCGACGCTACATTCAGCTATCCC	240		
Db	323	TGAGCAAGCAATTTTGAGAAATGGCTCTGTACTGCAACGCTACATTCAGTTATCCC	382		
Qy	241	TATATCCTTCGTTTCGGGTTTCTATGTACATGTGATGAGCCGCTGGTGGAGCCAGTA	300		
Db	383	TATATCCTTCGTTTCGGGTTTCTATGTACATGTGATGAGCCGCTGGTGGAGCCAGTA	442		
Qy	301	CGAGAACTTGGCCGTGACCGACCGCTTCATGATCCAGGTGCTTACGTTGCGTAGGGCAA	360		
Db	443	CGAGAACTTGGCCGTGACCGACCGCTTCATGATCCAGGTGCTTACGTTGCGTAGGGCAA	502		
Qy	361	GGATGAGGAAGCCGTTTGTGTCGGCGGAGAGATATCCGTTAGCCATCTTGGGCTAACT	420		
Db	503	GGAGCAAGCAAGGCGGTTGCTGCGGCGGACGCTATATCCGTACGGCATCTTGGGCTAACT	562		
Qy	421	GCTCATCTCGCGAGCATAGCAACTTGAGTCTACAAAGCGCTTTCACATCTTACCACT	480		

```

Db      563 GCTCATCTCGGAGCATCAGACACTCGCTACAGGCGTTTCCAGTCTTCAACACT 622
Qy      481 GGTCTTGCAGAGTTTATAGACCAATGGGAACTATAGAGTTGCGAAGTTGGGCTTACC 540
Db      623 GATCAAGCAGAGCTTTATAGACTCTCGGGGAATATAGAGTTGAGAGTTGGGCTTACC 682
Qy      541 ACACAACATCTTCTGGGTGCGGCTGGTGGTTGGTTGGCCAACTTGTCAATGAAGGCTTATCT 600
Db      683 ACACAAACATCTTCTGGGTGCGGCTGGTGGTTGGTTGGCCAACTTGTCAATGAAGGCTTATCT 742
Qy      601 TGAAGTGAATCCGGGACACCGTCTGCTCCAGAGCTGATGAATGAATGATGTACTTT 660
Db      743 TGAAGTGAATCAGGAGACCTGCTGCTCCAGAGCTGATGAATGAATGATGTACTTT 802
Qy      661 GCGTACTCAGTGTGAGACAGCTGTATGCTTACGATGATTAAGTATCCATTTGGTGTACAC 720
Db      803 GCGTACTCAGTGTGAGACCTATATGCTTATGATGATTAAGTATCCATTTGGTGTACAC 862
Qy      721 ACAGGTGATGACAGTGGAGATATACAGCTTTTCTGCTGATGATGAGGAGGAGT 780
Db      863 ACAGGTGATGACAGTGGAGGTTTACAGCTTTTCTGCTGATGATGAGGAGGAGT 922
Qy      781 TCTGAACCCCAAGAGACTACCCAGGCGCATGAGATGATGATGATGATGATGATGATGATGAT 840
Db      923 TCTGAACCCCAAGAGACTACCCAGGCGCATGAGATGATGATGATGATGATGATGATGATGAT 982
Qy      841 AATCTGCAATTTCTTATTTCTATGAGGCTGCTGATGATGATGATGATGATGATGATGATGAT 900
Db      983 AATCTGCAATTTCTTATTTCTATGAGGCTGCTGATGATGATGATGATGATGATGATGATGAT 1042
Qy      901 CTGCGGGAGGAGCATGATGATTTTGAAGTAACTAAGTAACTAAGTAACTAAGTAACTAAGTAACT 960
Db      1043 CTGCGGGAGGAGCATGATGATTTTGAAGTAACTAAGTAACTAAGTAACTAAGTAACTAAGTAACT 1102
Qy      961 GTCCCTGTGTCCGTGATGAGTACACAGAACTTGCCTCCATGAAACGTGACATGTA 1020
Db      1103 GTCCCTGTGTCTGTGATGAGTACACAGAACTTGCCTCCATGAAACGTGACATGTA 1162
Qy      1021 CTGGAACGAGGAGCGGCTTACGCGGCTTACACAGCTGCTTGTGCAAGTCTCCGCGCA 1080
Db      1163 CTGGAACGAGGAGCGGCTTACGCGGCTTACACAGCTGCTTGTGCAAGTCTCCGCGCA 1222
Qy      1081 TTCTTATGAGGCTTCACTTCAATGAGCTTAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1140
Db      1223 ATCTTATGAGGCTTCACTTCAATGAGCTTAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1274
Qy      1141 AAAAGAGAGCTGACACAGATAAGAAAGAGATGCTATAGCAGACCATAGGCTGCT 1200
Db      1275 -GAGGAGAGTACACTTATGACAAAGAAAGAAAGATGATTAAGGACGACCATTTGACGCTT 1333
Qy      1201 CTGGAACGAGCACCAGAAAGTACATCTTCTGTAAGACCTTAAAGACCAATCTAT 1260
Db      1334 CTGGAACGAGCACCAGAAAGTACATCTTCTGTAAGACCTTAAAGACCAATCTAT 1393
Qy      1261 GTGTCTTAAGA---ACCCCTCTCGAAGGCGAGTGAAGATGCAACGAGAAAGCA 1317
Db      1394 GTGTCTTAAGAAGACCCCTTCTCGAAGGCGAGTGAAGATGCAACGAGAAAGCA 1453
Qy      1318 GAAAGATGCTGGAATTTAAGGCTTGAATCTTGAATATGTTCCAAAGTTTAAAG 1377
Db      1454 GAAAGATG---ATGCTGGAAGAACAAAGAGACTCTTACAGACTATACAAAGTTTAAAG 1510
Qy      1378 GAGAGGCTCCATTGTGGGCCACAGGACCCAGACCA-----CCCTAC 1422
Db      1511 GCGCAAGTTACCGCCAGGCGCCACAGCCACTTGGCAGCCACTTCCATGTCTTCCCTCC 1570
Qy      1423 TGAAGAGTCAAGCAGCTC-----CAGTTGACACAGGTGATGGGCTTCCACAGATTA 1476
Db      1571 TCACAGTCAAGCAGCTCAGTGCCTTACTCAGACACAGGTGAGGCGCTTCCGACAGTA 1630
Qy      1477 CCAAGAAATCTGTCAATGAAAGAAAGAAAGTGTGAGTTTAACTTGAACATTTCCAGAGAG 1536

```

```

Db      1631 CCAAGAGCCTGTCAATGAAGAAAGAAAGAGTGAAGTTTAACTTGAACATTTCCAGAGTG 1690
Qy      1537 CCCACAGAACATTTGAACAGGCGGTTTGGACAGATGTCACAACTATACAGGCTCT 1596
Db      1691 CCTACGGAACA---TGAAACAACCCCATTTGACCAAGTG-----CACACAT 1735
Qy      1597 AATGAAGAGCATGACAGAGTCTTATCCCTACAGGAGTGAAGAGTGGACCAAACTGTCT 1656
Db      1736 CTTGAAGAGAGCTGAGAGATCTTATCTGTCAGATGAATGAAGCTGACACCAATCTGTTTC 1795
Qy      1657 CTATGATGATGCTTCAACAGCTGGCCCTGACTTGAAGATGTCACAGAGGAGCTGAC 1716
Db      1796 CTATGATGATGCTTCAACAGCTGGTTCTCACTGCACTGATGATG---CAGCAGGAGCTGAC 1854
Qy      1717 CCACTCAAGGACACAGAGGAGCAACCCAGAGGTGTTCCAGACAGCTTGAAGATG 1776
Db      1855 CCACTCAAGGAGCAC---AGCAGCAGACTCAAGAAATGTTCCATGAGCAGCTGCGCAGG 1913
Qy      1777 TAACTCAAGAACCAAGAGTAACTTAACTGATCTGCTGTAAGAACACCTGTAATTTAGATCTT 1836
Db      1914 TAACTCAAGAACCAAGAGTAACTGATCTGCTGTAAGAACACCTGATCTTCACTATCTT 1970
Qy      1837 T-CCCAACTGAAGATTTAATTAACCTGAATATTTCTTTTAAAGTGAAGAAAAA 1895
Db      1971 TCCCAAACTGACGAGTTAATA-----AATATTTCTTGAAGCAAAAAA 2024
Qy      1896 AAAAAAAAAA 1906
Db      2025 AAAAAAAAAA 2035

RESULT 4
AY357925
LOCUS      2187 bp mRNA linear PRI 26-AUG-2003
DEFINITION Macaca fascicularis beastrophin (VMD2) mRNA, complete cde.
ACCESSION AY357925
VERSION AY357925.1 GI:34013782
KEYWORDS
SOURCE      Macaca fascicularis (crab-eating macaque)
ORGANISM
REFERENCE
AUTHORS      Okamoto,H., Umeda,S., Suzuki,M.T., Yoshikawa,Y., Tanaka,Y. and
              Iwata,T.
TITLE      Molecular Cloning of VMD2 Gene from Cynomolgus Monkey (Macaca
              fascicularis)
JOURNAL      Unpublished
REFERENCE      2 (bases 1 to 2187)
AUTHORS      Okamoto,H., Umeda,S., Suzuki,M.T., Yoshikawa,Y., Tanaka,Y. and
              Iwata,T.
TITLE      Direct Submission
JOURNAL      Submitted (01-AUG-2003) National Tokyo Medical Center, National
              Institute of Sensory Organs, 2-5-1, Higasbisaoka, Meguro-ku, Tokyo
              152-0021, Japan
FEATURES
SOURCE
1..2187
    /organism="Macaca fascicularis"
    /mol_type="mRNA"
    /db_xref="taxon:9541"
1..2187
    /gene="VMD2"
67..1824
    /gene="VMD2"
    /codon_start=1
    /product="beastrophin"
    /protein_id="AA056049.1"
    /db_xref="GI:34013783"
    /translation="MTITVTSQVANNARLGSFSSLLCWRGSIYKLYGEFFILCY
    IIRFVRLATFEEQQLMFEKLTLYCDSYIQIIPISFLGFYVTLVYTRMNQYENLFW
    PDRIMSLVSGFVGEKDEGRLLRRTILRYANLGNVLIRSVTAIVYGRFPSSAHLVQA
    GFMTPEHKKLEKLSLPNNMFVPMWFWANLSMKAWLGRIIRDPILLQSLINENKTLR

```

10CGHLVAYDWISIPLYTVVAVTSFFLTCVGRQFLNPAKAYRGLHDLVVPE
 TFLQFFVYVWLVKVAEOLINPREGDDDDPENMTVDNLOVSLIADENHODIPREP
 DMYNPEPHPTAASOPRPARPMSSTRIISLKNKEMFQPNQEDKEDAHGIGR
 FLGASHHHPGANSRKLIMPRESLSLHGLEPKNKAKOVNGLKEDKAKLIV
 DAPKSAPIQPGPTSAVQPLSTPMEFPPEPEVLSKHSVTVGIDIKDLSLTVSSG
 AKNSSELISGSDGALMEHPESVSHVRKTVFENLDMBEIPENHLKEPBLESTNHAT
 LKDHVDPMVALENDRBAHS"

Query Match 45.6%; Score 873.8; DB 9; Length 2187;
 Best Local Similarity 77.2%; Pred. No. 2.5e-202;
 Matches 118; Conservative 0; Mismatches 292; Indels 39; Gaps 3;

QY 1 GTGCCAAGCCATGACTATCACTACACAAACAAAGTGGCCATGCGCGCTGGCTCGT 60
 DB 57 GTGCGTGGCCATGACCACTACCTACAGACCAAGTGGCTAAAGCCCGCTTAGGCTCTT 116
 QY 61 CTGCTCCCTCCTCCTGCTGGCGAGGACAGCATCTACAGCTGCTGTATGAGAAATTCCT 120
 DB 117 CTCCCGCTGCTGCTGTGGCGGAGACATCTACAGCTGCTGTATGCGAGTCTT 176
 QY 121 TGTCTTCAATCTCTACTATTCCATCCGTGGAAGCTACAGAAAGTCTCTCGAGTCA 180
 DB 177 CATCTTCTGCTCTGCTACTACATCATCCGCTTATTTATAGGCTGGCCCTCAGAGAA 236
 QY 181 TCAGCAGCTGTGTGTGAGAAAGCTGCTCTGTATCTGCGACAGCTACATTCAGCTATCC 240
 DB 237 GCAACAGCTGATGTGTGAGAAAGCTGACTGTATTCGACAGCTACATCAGCTATCC 296
 QY 241 TATATCTTCTGCTCTGGGTTTCTAATGTATCATTTGGGTGGACCGCTGTGAGGCACTA 300
 DB 297 CATTTCTTCTGCTGGCTTCTAGCTGACCTGGTGGTGGACCGCTGTGAGGCACTA 356
 QY 301 CGAGAACTTGGCGGCGCCGACCGCTCTCAATCCAGTGTCTAGCTGTGTGAGGCA 360
 DB 357 CGAGAACTTGGCGGCGCCGACCGCTCTCAATCCAGTGTGTGAGGCACTA 416
 QY 361 GGATGAGAAAGCCGCTTGTGGCGGAGCGCTCATCCGCTACGCTATCTGGGCAAGT 420
 DB 417 GGAACGAGAAAGCCGCTTGTGGCGGAGCGCTCATCCGCTACGCTATCTGGGCAAGT 476
 QY 421 GCTCATCTGCGGAGCATTCAGCACTCGGTCTACAAAGCGCTTCCCACTTTCACCACT 480
 DB 477 GCTCATCTTGGCGGAGCATTCAGCACTCGGTCTACAAAGCGCTTCCCACTTTCACCACT 536
 QY 481 GGTGCTACAGGTTTATGACCACTGAGGAACTATGAGAGTTGAGGCTTACC 540
 DB 537 GGTGCAAGAGCTTATGACCCCGGCAACCAAGAGTTGAGAGGCTTACC 596
 QY 541 ACACAAACATTTGGGTGGCTGGTGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 600
 DB 597 ACACAAACATTTGGGTGGCTGGTGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 656
 QY 601 TGGAGGTGGAATCCGGGACACCGCTGCTTCAGAGGCTGTATGATGAGGTGTGACTTT 660
 DB 657 TGGAGGTGGAATCCGGGACACCGCTGCTTCAGAGGCTGTGTATGATGAGGTGTGACTTT 716
 QY 661 GCGTACTCAGTGTGAGAGCTGTATGCTTCAAGCTGATATGATATCCCATTTGTTAC 720
 DB 717 GCGACTCAGTGTGAGAGCTGTATGCTTCAAGCTGATATGATATCCCATTTGTTAC 776
 QY 721 ACAGGTGTGAGAGCTGTATGAGCTTTTCTTGTGATGCTTGTGAGGAGGAGTT 780
 DB 777 ACAGGTGTGAGAGCTGTATGAGCTTTTCTTGTGATGCTTGTGAGGAGGAGTT 836
 QY 781 TCTGAACCAACAAAGAGCTACAGGCTATGAGATCTGTGTGCTGTCTTAC 840
 DB 837 TCTGAACCAACAAAGGCTTACCTTCCATGAGCTGTGAGCTGTGTGCTGTCTTAC 896
 QY 841 AATCTGCAATTTCTTATCTACATGAGCTGTGAGAGTGTGAGAGCAAGCTATGAACC 900
 DB 897 AATCTGCAATTTCTTATCTTATGAGTGTGAGTGTGAGAGTGTGAGAGCAAGCTATGAACC 956

QY 901 CTTCGGGAGAGACATGATGATTTTGAAGCTAATGATCATTTGACAGAAACCTGCAGCT 960
 DB 957 GTTTGGAAGAGACATGATGATTTTGAAGCACTGATGATTTGACAGAAATTTGACAGG 1016
 QY 961 GTTCCTGTGTTCCGTGATGAGATGACACAGAACTTGCCTCCATGAGAAAGTACATGT 1020
 DB 1017 GTTCCTGTGCTGTGTGATGATGATGACACAGAACTTGCCTCCATGAGAAAGTACATGT 1076
 QY 1021 CTGGAACAGAGCAGCGCTCAGCGCCCTACACAGCTGCTTTCAGAGTCTTGCAGGCA 1080
 DB 1077 CTGGAATGAGCTGTAGCAACACCCCTTACACAGCTGCTTTCAGAGTCTTGCAGGCA 1136
 QY 1081 TTCTTCAATGAGCTCAGCTTCAACATCAGCTTAAAGAAAGAGCTTGAAGTTCAGC 1140
 DB 1137 CTCTTATGAGCTCAGCTTCAACATCAGCTTAAAGAAAGAGCTTGAAGTTCAGC 1196
 QY 1141 AAAAGAGAGCTGACACGAGTAAAGAAAGAGTGTATGACAGACCATAGCTGCTT 1200
 DB 1197 CAATCAGAGG-----ACAAAGAGAGCTCAGCTGATCATTTGGCGCTT 1244
 QY 1201 CTTAGAGCTGCAACCCAAACCTACCATCTTCCCTTGAAGACTTAAAGACCAATAT 1260
 DB 1245 CTAGAGGCTGCAATCCCATGATCATCTTCCGAGGCAACTCAAGAACCAATCT 1304
 QY 1261 GTGTTCTAAG-----ACCCCTCTGAGAGGCACTGTAAGATGCCAACAGAAAA 1314
 DB 1305 GTGCGCCAAAGAGGAATCCCTTCTCAGAGGCGCTGCCAAACCAAGGAGTCA 1364
 QY 1315 CCAGAAAG-----ATGCTGGAATTTAAGGCTGTGAGCTTTT 1353
 DB 1365 ACAGAAAGTTAGGGGCTGGAAGACAAAGAGCTTGAAGCTTGAAGTCTT 1424
 QY 1354 GAAATGTGTTCCAGGTTTAAAGAGAGAGCTCCCATTTGTGGCCACAGGACCCAGAG 1413
 DB 1425 CAAGTGTGCCCACTGATCAGAGGCGCAGCTACTACAGTCCCAAGAGAGCCCTCAG 1484
 QY 1414 CCACCTTAC 1422
 DB 1485 CCCCACTCC 1493

RESULT 5
 AF073501 2171 bp mRNA linear PRI 02-SEP-1998
 LOCUS
 DEFINITION
 Homo sapiens vitelliform macular dystrophy protein (VMD2) mRNA,
 complete cds.
 ACCESSION
 AF073501
 VERSION
 AF073501.1 GI:3511241
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)

REFERENCE
 1 (bases 1 to 2171)
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 Stohr, H., Marguardt, A., Rivera, A., Cooper, P. R., Nowak, M. J.,
 Shows, T. B., Gerhard, D. S., and Weber, B. H.,
 A gene map of the Best's vitelliform macular dystrophy region in
 chromosome 11q12-q13.1
 Genome Res. 8 (1), 48-56 (1998)
 MEDLINE
 JOURNAL
 PUBMED
 9445467
 REFERENCE
 2 (bases 1 to 2171)
 Marguardt, A., Stohr, H., Paesmore, L., Kraemer, F., Rivera, A. and
 Weber, B. H. F.,
 Direct Substitution
 Submitted (22-JUN-1998) Human Genetics, University, Biozentrum, Am
 Hubland, Wuerzburg 97074, Germany
 Location/Qualifiers
 1..2171
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="11"

FEATURES
 source

TITLE Identification of the gene responsible for Best macular dystrophy
JOURNAL Nat. Genet. 19 (3), 241-247 (1998)
MEDLINE 98324772
PUBMED 9662395
REFERENCE 2 (bases 1 to 2210)
AUTHORS Petrukhin, K.
TITLE Direct Submission
Submitted (03-APR-1998) Human Genetics, Merck Research
Laboratories, West Point, PA 19486, USA
FEATURES
source
1..2210
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q13"
1..2210
/gene="VMD2"
105..1862
/gene="VMD2"
/note="Best's macular dystrophy gene; BMD; alternatively
splice isoform"
/product="bestrophin"
/codon_start=1
/product="bestrophin"
/protein_id="AA064343.1"
/db_xref="GI:335159"
/translation="MTITYSQVANARLGSFSRLILCWGRSGIYKLKLYGEFLFLICY
ILREIYRLALTEBOQLEKLTLYCDVSYQLIPISFVIGFYTVLVTVTMNOYENLPA
PRMLSLVSGFVGEKDEQGLRLRLIRYANLGNVLILRSVSTAYKRPFAHQYQA
GMPTAEHKEKLSLPHNFWPMVPMWPMANISMKMILGGRIRDPILLOSLENMTLR
TOCGHLXAYDWISILPLYOVTVAVSPEFLICVGRPLNPKAYREHELDLVVPR
TLOPFYFYGVKVAEOLINPFGDDDDDEFETWIVDRNLQVSLAVDEHQLDPRMEP
DMYMKREPQRYTAAQAFRPAFSGSTFENISLNKEMEQPOVEBEDNAHAIIGR
FGLQSHDHPHPRANSRTKLMIPKRESILHBLPNHNAKAKONVGOEDNDKMLKAV
DAKSAPLYORPGYSAPOPLSPPTMFEPLSPASPSLHSGTIGDTDKSLKTVSSG
AKSFPELLSESDGALMEHPEVSVQRKTVRENLTDMPETIPENHLKEPLSGSPNINHT
LKHNDPYVALENRDEAHS"

ORIGIN

Query Match 45.5%; Score 871.2; DB 9; Length 2210;
Best Local Similarity 77.1%; Pred. No. 1.1e-201;
Matches 1116; Conservative 0; Mismatches 293; Indels 39; Gaps 3;

```
QY 2 TCCCAAGCCATGACATATCACTACCAAAAGTAGCCATGCCCGCTGGTTGCTTC 61
DB 96 TCCCTGGCCATGACATATCACTACCAAGTAGCCATGCCCGCTGGTTGCTTC 155
QY 62 TCGTCCCTCTCTCTGTGTGCGAGGAGCATCTTACAAGCTGTGTATGAGAATTGCTT 121
DB 156 TCCCGCTGTGTGTGTGTGCGAGGAGCATCTTACAAGCTGTGTATGAGAATTGCTT 215
QY 122 GTCTTCATATTCCTCTACTATTCATCCGTGAGACTCTACAGAAATGTTCTCTGAGTAT 181
DB 216 ATCTTCCGTCTGTCTACTACATATCCGCTTATTTATAGGCTGGCCCTCAAGAA 275
QY 182 CAGAGCTGTGTGGAAGACTGTCTGTATCTGCACAGACTTACATTCAGCTCATCCCT 241
DB 276 CAACAGCTGATGTTTGAAGAACTGACTGTATTTGGACAGCTACATCCAGCTCATCC 335
QY 242 ATATCTCTGTTCTGGGTTTCTAATGTATCAATGGTGTAGCCGCTGGTGAACCAAGTAC 301
DB 336 ATTTCTCTTGCTGGGCTTCTAATGTATCAATGGTGTAGCCGCTGGTGAACCAAGTAC 395
QY 302 GAGAACTTGGCCGCGCCGACCGCTCATGATCCAGGTGTCTAGCTTCTGTGAGGGCAAG 361
DB 396 GAGAACTTGGCCGCGCGCCGACCGCTCATGATCCAGGTGTCTGTGAGGGCAAG 455
QY 362 GATGAGAAAGCGGTTTGTCTGTGCGCGACGCTCATCCGCTACCGCAATCCCTGGGCAAGTG 421
DB 456 GACGAGAAAGCGGCTGTGCTGTGCGCGACGCTCATCCGCTACCGCAATCCCTGGGCAAGTG 515
QY 422 CTCATCTGTGCGAGCATCAGCACTGAGTCTAACAGCGCTTCCACTTTTACCAACCTTG 481
|||||
```

```
DB 516 CTCATCTGTGCGAGCATCAGCACTGAGTCTAACAGCGCTTCCCAAGCGCCACAGCACTG 575
QY 482 GTGCTAGCAGGTTTTATGACCCATAGGGGAACATAGAAGTTTGCAGAAATTGGGCTTACCA 541
DB 576 GTGCAAGCAGGCTTTATGATCTCCGCGACAAACAGAGATTTGGAAGAACTGACCTTACCA 635
QY 542 CACAACATCTTCTGGGTGCCCTGGGTGTGGTTTGGCCAATTGTCAATGAGAGCTTACTT 601
DB 636 CACAACATGTTCTGGGTGCCCTGGGTGTGGTTTGGCCAATTGTCAATGAGAGCTTACTT 695
QY 602 GAGAGTGAATCCGGGACACCGCTCTGCTCCAGAGCTGATGATGATGAGTGTACTTTG 661
DB 696 GAGAGTGAATCCGGGACACCGCTCTGCTCCAGAGCTGATGATGATGAGTGTACTTTG 755
QY 662 CGTACTAGTGTGACAGCTGTATGCTTACAGCTGTATGATTAAGTATCCCATTTGGTACACA 721
DB 756 CGTACTAGTGTGACAGCTGTATGCTTACAGCTGTATGATTAAGTATCCCATTTGGTACACA 815
QY 722 CAGGTGTGACAGCTGTATGCTTACAGCTGTATGATTAAGTATCCCATTTGGTACACA 781
DB 816 CAGGTGTGACAGCTGTATGCTTACAGCTGTATGATTAAGTATCCCATTTGGTACACA 875
QY 782 CTGAACCAAAACAAGACTTACCCAGGCAATGAGATCTGGTTGTGCTGTTCACA 841
DB 876 CTGAACCAAAACAAGACTTACCCAGGCAATGAGATCTGGTTGTGCTGTTCACA 935
QY 842 ATCTCTGAATTTCTTATTTTCAATGAGCTGTGAAAGTGTGAGAAAGCTCATCAACCC 901
DB 936 TTTCTCTGAGTCTTCTTCTATGATGCTGTGCTGGAAGTGTGAGACAGCTCATCAACCC 995
QY 902 TTGGGAGAGACAGATGATATTTTGAAGTAACTGAGATCATTTGACAGAAACCTGAGGTG 961
DB 996 TTTGAGAGAGATGATATATTTTGAAGTAACTGAGATCATTTGACAGAAATTTGACAGGTG 1055
QY 962 TCCCTGTGTTCCGTGATGGAATGACAGAACTTCCCTCCATGGAAGTGAACATGTAC 1021
DB 1056 TCCCTGTGTTCCGTGATGGAATGACAGAACTTCCCTCCATGGAAGTGAACATGTAC 1115
QY 1022 TGAACAGAGAGAGGCTTCAAGCGGCTTACACAGCTGCTTGTCCAGGTTCTGCGCGCAT 1081
DB 1116 TGAATTAACCGGACCAAGCGGCTTACACAGCTGCTTGTCCGCGCATGCTGAGGC 1175
QY 1082 TCTTCATGAGGCTTCAAGCTTCAATCACTGAGCTTAAAGAAAGAACTTAAAGCTTGTGCA 1141
DB 1176 TCTTCATGAGGCTTCAAGCTTCAATCACTGAGCTTAAAGAAAGAACTTAAAGCTTGTGCA 1235
QY 1142 AAAGAGAGGCTGACACGGAATPAAGAAAGAGTGTATGACAGCACCATAGCTGCTTC 1201
DB 1236 AATCAGAGAGAGAGAGAGATGCTCAGGCTG-----GATCATTTGGCGGCTTC 1283
QY 1202 TTAGAGCTGCAACCAAAATACCATCTTCCCTTGAAGCTTAAAGCAAAATTTG 1261
DB 1284 CTAGGCTGTGAGTCCCATGATGATCAATCTTCCAGGCGAAATCTAAGAACTTAACTG 1343
QY 1262 TGTCTTAAGA-----ACCCCTCTCTGAGAGGCAAGTGAAGATGCAACCAAAATG 1315
DB 1344 TGGCCCAAGAGGAATCTTCTTCAAGAGGCTGTGCCCAAAACCAAGAGCAAGCAAA 1403
QY 1316 CAGAAAG-----ATGCTGGAATAATTTAAAGGTCTGAGCTTCTTG 1354
DB 1404 CAGAAAGCTTAAAGGCGCAGAGAAACAAACAAAGGCTTGAAGGCTGTGAGCGCTTTC 1463
QY 1355 AAATGTGTTCAGAGTTTAAAGAGAGAGGCTCCATTTGTGAGCCCAAGGCAACCAAGCAG 1414
DB 1464 AAGTGTGCGCACTGTATCAGAGGCGAGGTACTTACAGTGTCCCAAGAGCGCCCTCAGC 1523
QY 1415 CACCTTAC 1422
DB 1524 CCACTTC 1531
RESULT 7
BD136720
```

LOCUS BD136720 2229 bp DNA linear PAT 18-SEP-2002
DEFINITION Best's macular dystrophy gene.
ACCESSION BD136720
VERSION BD136720.1 GI:23231665
KEYWORDS JP 2002504559-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2229)
Petrukhin, K., Caskey, T.C., Metzker, M. and Wadelius, C.
Best's macular dystrophy gene
Patent: JP 2002504559-A 2 12-FEB-2002;
MERCK & CO INC, CLASS MADELIUS
COMMENT OS Homo sapiens (human)
PN JP 2002504559-A/2
PD 12-FEB-2002
PF 22-FEB-1999 JP 2000533447
PR 25-FEB-1998 US 60/075941, 18-DEC-1998 US 60/112926 PI
KONSTANTIN PETRUKHIN, THOMAS C CASKEY, MICHAEL METZKER, CLAES PI
MADELIUS
PC C07K16/18, C07K14/47, C12N5/10, C12N15/09, C12P19/34, C12Q1/68// PC
C12P21/08,
PC C12N5/00, C12N15/00
CC Best's macular dystrophy gene
FH Key 1.2229 Location/Qualifiers
FT source 1.2229 /organism="Homo sapiens (human)".
FEATURES
source 1.2229
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 45.5%; Score 871.2; DB 6; Length 2229;
Best Local Similarity 77.1%; Pred. No. 1.1e-201;
Matches 1116; Conservative 0; Mismatches 293; Indels 39; Gaps 3;
QY 2 TGCCAAAGCATGACTATCATCTACCAACAAAGTAGCCCATGCCGCTCGGTTGCTTC 61
DB |||||
QY 96 TGCCGCGAGTACATCATCTTACCAAGCCAGTGCTATGCCGCTTAGAGCTCTTC 155
DB |||||
QY 62 TCGTCCCTCTCTGCTGTGGCGAGCATCTTACAAAGCTGCTGTATGAGAACTTCTT 121
DB |||||
QY 156 TCCGCGCTGCTGTGTGGCGAGCATCTTACAAAGCTGCTGTATGAGAACTTCTT 215
DB |||||
QY 122 GTCTTCATATCTCTACTATTTCCATCCGTGGAATCTTACAGATGAT 181
DB |||||
QY 216 ATCTTCTGCTGTGCTACTATCATCTCGCTTTATTTATAGGCTGGCCCTACGGA 275
DB |||||
QY 182 CAGAGCTGTTGTTAGAAAGCTGCTGTACTGCGACAGCTACATTCACTCATCCCT 241
DB |||||
QY 276 CAAAGCTGATGTTGAGAACTGACTGTATTGCGACACTACATCACTCATCCCT 335
DB |||||
QY 242 ATATCTTCTGTTGGGTTTCTATGTTATCTTGTGTGAGCCGCTGTGAGCCAGTAC 301
DB |||||
QY 336 ATTTCTTCTGCTGTGGCTTTTACGTGACGCTGTGACCCGCTGTGGAACCAAGTAC 395
DB |||||
QY 302 GAGAACTTGCCTGAGCCGCAAGCTCATGATCCAGGTGTCTAGCTTCTGTGAGGCAAG 361
DB |||||
QY 396 GAGAACTGCTGCTGAGCCGCAAGCTCATGAGCTGTGTGAGGCTTGTGAGGCAAG 455
DB |||||
QY 362 GATGAGAAAGCCGTTTGTGCGGCGACGCTCATCGCTACGCCATCTTGAGGCAATG 421
DB |||||
QY 456 GACGAGAAAGCCGCTGCTGCGGCGACGCTCATCGCTACGCCATCTTGAGGCAATG 515
DB |||||
QY 422 CTATCTCTGCGCAGATAGACCTCTGCTTACAAAGCTTTCCTTCCATCTTCACTG 481
DB |||||
QY 516 CTCTATCTGCTGCGCAGCTAGACCGCACTTACAAAGCTTTCCTTCCATGAGCCAG 575
DB |||||
QY 482 GTGCTAGAGGTTTATGACCATGAGGAGATAGAGAGGTTGAGAGGTTGAGGCTTACA 541
DB |||||

DB 576 GTGCAAGAGGCTTTATGACTCCGCGAGAACACAAAGAGTTGAGAACTGAGCTTACCA 635
QY CACAACATCTCTGGGCGCTCGGCTGTGCTTTCGCACTTGTAAAGAGCCATCTT 601
DB |||||
QY 636 CACAACATCTCTGGGCGCTCGGCTGTGCTTTCGCACTTGTAAAGAGCCATCTT 695
DB |||||
QY 602 GAGAGTGAATCCCGGAGCACCGCTGTCTCCAGAGCCGTATGAAAGAGTGTACTTTG 661
DB |||||
QY 696 GAGAGTGAATCCCGGAGCACCGCTGTCTCCAGAGCCGTATGAAAGAGTGTACTTTG 755
DB |||||
QY 662 GTTACTAGTGTGACAGCTGTATGCTTACGACTGATAGTATCCATTTGGTGTACCA 721
DB |||||
QY 756 CGTACTAGTGTGACAGCTGTATGCTTACGACTGATAGTATCCATTTGGTGTATCA 815
DB |||||
QY 722 CAGGTGTGACAGTGTATGAGCTTCTTCTGCTGCTGTATGAGGAGGAGTGT 781
DB |||||
QY 816 CAGGTGTGACAGTGTGAGGCTGTATGAGCTTCTTCTGCTGCTGTATGAGGAGGAGTGT 875
DB |||||
QY 782 CTGAACCCAAACAGAGCTACCCAGGCGATGAGATGATCTGTGTGCTGTCTTACA 841
DB |||||
QY 876 CTGAACCCAGGCGAGGCTTACCTGCGCATGAGCTGACCTGTTGCTTTCACG 935
DB |||||
QY 842 ATCTGCAATTTCTTATTTCTACATGAGGCTGCTGTAAGGTGAGAAACAGCTTCAACCC 901
DB |||||
QY 936 TTCTGCGAGTTCTTCTTATGTTGCTGCTGTAAGGTGAGAGAGCTTATCAACCC 995
DB |||||
QY 902 TTGCGGAGAGCATGATGATTTTGAAGCTTAACTGATCATTTGACAAACCTTCAAGTG 961
DB |||||
QY 996 TTTGAGAGAGTGTATGATTTTGAAGCAACCTGATTTGTCACAGAAATTTTCAAGTG 1055
DB |||||
QY 962 TCCGCTGTGCTGCTGATGAGGATGACACCAAGATTTGCTCCATGAAAGCTGATGAC 1021
DB |||||
QY 1056 TCCGCTGTGCTGCTGATGAGATGACACCAAGATTTGCTCCGATGAGAGCCGAGTATGAC 1115
DB |||||
QY 1022 TGAAGAGAGGAGGCTTACAGCCGCTTACACAGCTGCTTGTGCGAGGCTCTCGCGGAT 1081
DB |||||
QY 1116 TGAATTAAGCCCGAGCCACAGCCCTTACACAGCTGCTTGTGCGAGGCTCTCGCGGAT 1175
DB |||||
QY 1082 TCTTTCATGGGCTTCACTTCAATCATGAGCTTAAAGAAAGACTTGTGATTTGTGCA 1141
DB |||||
QY 1176 TCTTTCATGGGCTTCACTTCAATCATGAGCTTAAAGAAAGACTTGTGATTTGTGCA 1235
DB |||||
QY 1142 AAAGAGAGGCTGACAGGATTAAGAAAGAGTGTATGACAGCATCATGAGCTTGTG 1201
DB |||||
QY 1236 AATGAGAGAGGAGGAGTGTCTACGCTG-----GATATTTGGCGCTTC 1283
DB |||||
QY 1202 TTAGAGTGCACAAACCAAACTACCATCTTCTTGAAGAATTAAAGACCAAACTATG 1261
DB |||||
QY 1284 CTAGGCTGCGAGTCCATGATCACCATCTTCCAGGGCAAACTCAAGGACCAAACTACTG 1343
DB |||||
QY 1262 TGTTCCTAAG-----ACCCCTCTCTGAGAGGCCAGTGTAAAGATGCCAACAGAAAAC 1315
DB |||||
QY 1344 TGCGCCAAAGAGGAATCCCTTCTCCAGAGGCGCTGCGCAAAAACCAAGAGCGCAAA 1403
DB |||||
QY 1316 CAGAAAG-----ATGCTGAAATTTAAGGGCTGTGACTTTG 1354
DB |||||
QY 1404 CAGAACTGTAGGGCCAGAGAAAGACAAAGAGCTGTGAAGCTTAAAGCTGTGAAGCTTTC 1463
DB |||||
QY 1355 AATATGTTCCAGATTTTAAAGAGAGGCTTCCATTTGTGCGCAAGGACCAAGCAGC 1414
DB |||||
QY 1464 AAGCTGAGCCACGTATCAGAGGCGAAGGCTTAAAGTGTGCGCAAGGACCAAGCAGC 1523
DB |||||
QY 1415 CACCTTAC 1422
DB |||||
QY 1524 CCCACTCC 1531
DB |||||
RESULT 8
AX745964 1758 bp DNA linear PAT 14-MAY-2003
LOCUS AX745964
DEFINITION Sequence 1 from Patent WO03030922.
ACCESSION AX745964
VERSION AX745964.1 GI:30724619

KEYWORDS Homo sapiens (human)

SOURCE

ORGANISM Homo sapiens

REFERENCE 1 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. Steuernagel, A., Bruenner, G., Fritsch, R., Eulenberger, K. and Closek, T.

TITLE Bestrophin and bestrophin homologous proteins involved in the regulation of energy homeostasis

JOURNAL Patent: WO 03030922-A.1 17-APR-2003; Develoen Aktiengesellschaft fuer Entwicklungsbiologische Forschung (DE)

FEATURES

source Location/Qualifiers

1..1758

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

ORIGIN

Query Match 45.3%; Score 867; DB 6; Length 1758;

Best Local Similarity 77.1%; Pred. No. 1.1e-200;

Matches 110; Conservative 0; Mismatches 290; Indels 39; Gaps 3;

QY 11 ATGACTATCACTACACAAACAAGTACCAATGCCCGCTGCTTCTGCTCCCTC 70

Db 1 ATGACCATCACTTACACAGCAAGTGGCTTAATGCCGCTTCTGCTCCGCTG 60

QY 71 CTCCTGTGCTGGCAGGACATCTACAAAGTCTGTATGAGAAATTCCTGTCTCA 130

Db 61 CTCCTGTGCTGGCAGGACATCTACAAAGTCTGTATGAGAAATTCCTGTCTCA 120

QY 131 TTCCTCACTATTCCTCCGTGACCTCTACAGAAATGTTCTCGATGATCAGAGCTG 190

Db 121 CTCGTGCTACATCAATCCGCTTATTTATAGAGCTGGCCCTCAGGAAACACAGCTG 180

QY 191 TTGTTGAGAAAGTGGCTCTGTACTGACAGCTACATTCAGTCACTTCCCTTC 250

Db 181 ATGTTGAGAAAGTGGCTCTGTATTGTCGACACTACACTACCTATCCCTATTCCTTC 240

QY 251 GTTCTGGGTTCTATGTTTACATTGTGTGAGCCGCTGTGTAGAGCAGTACAGAACTTG 310

Db 241 GTGCTGGGTTCTATGCTGACGCTGTGTGTGACCCCTGTGTGAAACAGTACAGAACTTG 300

QY 311 CGGTGGCCGACCGGCTCATGATCCAGGATCTAGCTTCGTGAGAGGAGAAAGTGAAGAA 370

Db 301 CGGTGGCCGACCGGCTCATGATCCAGGATCTAGCTTCGTGAGAGGAGAAAGTGAAGAA 360

QY 371 GGCCGCTTGTGCGGCGGACGCTCATCCGCTACGCGATCCGAGCAGAGTCTCATCTG 430

Db 361 GGCCGCTTGTGCGGCGGACGCTCATCCGCTACGCGATCCGAGCAGAGTCTCATCTG 420

QY 431 CGAGCATCAGACCTCGGTCTACAAAGGCTTCCACTCTTTCACCACTGTGTGTCGA 490

Db 421 CGAGCATCAGACCTCGGTCTACAAAGGCTTCCACTCTTTCACCACTGTGTGTCGA 480

QY 491 GGTTTTATGACCAATGGGAAACATAGCAAGTTGAGAAAGTTGGGCTTACCAACACA 550

Db 481 GGTTTTATGACCTCGGGAACAACAAGTTGAGAAAGTTGGGCTTACCAACACA 540

QY 551 TTCTGGGAGCCCTGGGTGTGTTGCCAATCTGTCAATGAAGCCATATCTTGGAGTGA 610

Db 541 TTCTGGGAGCCCTGGGTGTGTTGCCAATCTGTCAATGAAGCCATATCTTGGAGTGA 600

QY 611 ATCCGGGACACCGTCTGCTCAAGAGCTGATGAATGAGGTGTGTTCTTGGGATCAG 670

Db 601 ATCCGGGACACCGTCTGCTCAAGAGCTGATGAATGAGGTGTGTTCTTGGGATCAG 660

QY 671 TGTGAGAGCTGATGCTGACATGATGATATCCATGTGTGACACAGGTGTG 730

Db 661 TGTGAGAGCTGATGCTGACATGATGATATCCATGTGTGATACAGGTGTG 720

QY 731 ACAATGGAGATATACAGCTTTTCTTGTGATGCTGATGCGGAGGAGATTTCTGAACCA 790

Db 721 ACTGTGGCGGTGTACAGCTTCTTCTGACTGTGTAGTGGCGGACAGTTCTGAACCA 780

QY 791 AACAGGACTACCCAGGCGCATGATGATGTGTGTGCTGTCTTACAACTCTGCAA 850

Db 781 GCCAAGGCTTACCCCTGAGCATGAGCTGTGACTCTTGTGCCCCCTTCAAGTCTCTGAG 840

QY 851 TTCTTATTTCTACATGAGGTGAGCTGAGAGTGGCACAACAGCTCATCAACCCCTCGGGAG 910

Db 841 TTCTTCTTCTATGTTGTGTGCTGATGATGAGAGCAAGCTCATCAACCCCTTGGAGAG 900

QY 911 GACGATGATGATTTTGAAGACTTAATGATATGACAGAAACCTGACAGTGTCTCTGTTG 970

Db 901 GATGATGATGATTTTGAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATG 960

QY 971 TCCGTGATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1030

Db 961 GCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020

QY 1031 GCAGGCTTCAAGCCGCTTACAGAGCTGCTTGTGCAAGGCTGTGCGGCAATTCCTCATG 1090

Db 1021 CCCAGGCAAGCCGCTTACAGAGCTGCTTGTGCAAGGCTGTGCGGCAATTCCTCATG 1080

QY 1091 GGCTCCACCTTCAACATGAGCCCTTAAAGAAAGAAAGACTTGTGCTTGTGCTTGTGCTTGTG 1150

Db 1081 GGCTCCACCTTCAACATGAGCCCTTAAAGAAAGAAAGACTTGTGCTTGTGCTTGTGCTTGTG 1140

QY 1151 GCTGACAGGATTAAGAAAGAGTGTGCTTATGACAGACCATGAGCTGTCTTGTAGACTG 1210

Db 1141 G-----ACGAGAGAGATGTCTACGCTGGGATATGTGCTTGTAGACTG 1188

QY 1211 CAACCCAAAACATCATCTTCCCTTGAAGACTTAAAGCAAACTATGTGTCTTAAAG 1270

Db 1189 CAGTCCCATGATCATCTTCCCTTGAAGACTTAAAGCAAACTATGTGTCTTAAAG 1248

QY 1271 A-----ACCCCTCTCTGAAAGGCAAGTGTAGATGATGATGATGATGATGATGATGATG 1322

Db 1249 AGGAAATCCCTTCTCAGAGAGGCTGCTGCAAAACCAAGGAGCCCAAGAACTT 1308

QY 1323 -----ATGTGGAATTTAAAGGTTCTGACTTCTGAAATGTGTT 1363

Db 1309 AGGAGGCAAGAAACAAAGGCTTGAAGCTTAAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAG 1368

QY 1364 CCAAGGTTTAAAGAGAGGCTCCCATTTGTGAGCCCAAGGACCAAGGACCACTTAC 1422

Db 1369 CCAGTGTATCAGAGGCAAGGCTTACAGTGTGCTTCAAGGACCAAGGACCACTTAC 1427

RESULT 9

AY515704 1758 bp mRNA linear PRI 28-JAN-2004

LOCUS

DEFINITION

AY515704 Homo sapiens bestrophin 1 mRNA, complete cds.

ACCESSION

AY515704.1 GI:41216872

VERSION

KEYWORDS

SOURCE

ORGANISM Homo sapiens (human)

REFERENCE

1 Homo sapiens

2 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

3 Sun, H., Tsunenari, T., Yau, K. W., and Nathans, J.

4 The vitelliform macular dystrophy protein defines a new family of chloride channels

5 Proc. Natl. Acad. Sci. U.S.A. 99 (6), 4008-4013 (2002)

6 11904445

7 2 (bases 1 to 1758)

8 Tsunenari, T., Sun, H., Williams, J., Cahill, H., Smallwood, P., Yau, K. W., and Nathans, J.

9 Structure-function analysis of the bestrophin family of anion channels

10 JOURNAL

11 J. Biol. Chem. 278 (42), 41114-41125 (2003)

12 22917481

JOURNAL thereof
 Patent: WO 02068579-A 15567 06-SEP-2002;
 PE Corporation (NY) (US)
 FEATURES Location/Qualifiers
 source 1. 2170
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Query Match 44.8%; Score 859.2; DB 6; Length 2170;
 Best Local Similarity 77.0%; Pred. No. 9.3e-199;
 Matches 1115; Conservative 0; Mismatches 293; Indels 40; Gaps 4;

2 TGGCAAGCCATGATCTATCACTTACCAACAAGTGGCCATGCGCGCTGGTCTG 61
 80 TGGCTGGCCATGACATCTTACACAGCCAGG-GGCTAATGCCGCTTAGGCTCTTC 138
 62 TCGTCCCTCTCTGCTGCTGGGAGGAGCATCTACAGCTGCTGTATGAGAAATTCCT 121
 139 TCCCGCTGCTGCTGCTGGGAGGAGCATCTACAGCTGCTATATGCGAGTTCTTA 198
 122 GTCTCATATTCCTCTACTATTCCATCCGAGCATCTACAGATGTTCTTCAGATGAT 181
 199 ATCTTCTGCTGCTGCTACTACATCTCCGCTTATTTATGAGCTGCTGCTCAGGAGAA 258
 182 CAGACGCTGTTTGGAGAGCTGCTCTGTAAGTGGAGCTTACATTCAGTCACTCCCT 241
 259 CACAGCTGATGTTGAGAACTGACTGTGATGAGACGCTACATCCAGTCACTCCCT 318
 242 ATATCCTGCTGCTGGGTTTCTATATGATGTTGTTGTTGAGCGCTGGTGGAGCAATAC 301
 319 ATTTCTTCTGCTGGGCTTCTACGTTGAGGCTGCTGAGACCCGCTGGTGGAGCAATAC 378
 302 GAGAACTTGGCGTGGCCGAGCTCTCATGATCCAGTGTCTAGCTGTGGAGGAGGAG 361
 379 GAGAACCTGGCGTGGCCGAGCTCTCATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 438
 362 GATGAGGAGGCGCTTGTGCTGGGAGGAGCTGATCCGCTACGCCATCTGGGAGCAATG 421
 439 GACGAGGAGGCGCTGCTGCTGGGAGGAGCTGATCCGCTACGCCATCTGGGAGCAATG 498
 422 CTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 481
 499 CTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 558
 482 GTGCTGAGGCTTATGAGCCATGAGGAGCAATGAGCTGAGGAGGCTGAGCA 541
 559 GTGCAAGGAGCTTATGACTCCGAGAGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAG 618
 542 CACAAACATTTCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 601
 619 CACAAACATTTCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 678
 602 GAGAGTGAATCCGAGCAACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 661
 679 GAGAGTGAATCCGAGCAACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 738
 662 CGTACTGAGTGGAGCAAGCTGTATGCTGAGCTGAGTAAAGTATCCGATTTGGTGTACACA 721
 739 CGTACTGAGTGGAGCAAGCTGTATGCTGAGCTGAGTAAAGTATCCGATTTGGTGTACACA 798
 722 CAGGTGAGCAAGGAGTAAACAGCTTTTCTGCTGAGCTTGTGAGTGGAGGAGGAGGAGT 781
 799 CAGGTGAGCAAGGAGTAAACAGCTTTTCTGCTGAGCTTGTGAGTGGAGGAGGAGGAGT 858
 782 CTGAACCCAAACAGAGCTACCCAGGAGCATGAGTGGTGTGCTGCTGCTGCTGCTGCTG 841
 859 CTGAACCCAAACAGAGCTACCCAGGAGCATGAGTGGTGTGCTGCTGCTGCTGCTGCTG 918
 842 ATCTGAGTATTTATTTATGAGGCTGCTGAGTGGAGGAGGAGGAGGAGGAGGAGGAGG 901
 919 TTTCTGAGTATTTATTTATGAGGCTGCTGAGTGGAGGAGGAGGAGGAGGAGGAGGAGG 978

QY 902 TTCCGGGAGGAGCATGATGATTTTGAAGTACTAGTATCATTTGACAGAAACCTGACAGT 961
 DB 979 TTGAGAGAGATGATGATGATTTTGAAGCAACATGAGTGTTCGACAGAAATTTGACAGT 1038
 QY 962 TTCCCTGTTGCTCCGTTGATGAGGATGACCAAGCTTGGCTCCCATGAGAACTGACATGAT 1021
 DB 1039 TTCCCTGTTGCTCCGTTGATGAGGATGACCAAGCTTGGCTCCCATGAGAACTGACATGAT 1098
 QY 1022 TGAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1081
 DB 1099 TGAATTAAGCCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1158
 QY 1082 TCCCTTATGAGGCTCACTTCAATCATGAGCTTAAAGAGAGGAGGAGGAGGAGGAGGAGG 1141
 DB 1159 TCCCTTATGAGGCTCACTTCAATCATGAGCTTAAAGAGAGGAGGAGGAGGAGGAGGAGG 1218
 QY 1142 AAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1201
 DB 1219 AATCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1266
 QY 1202 TTGAGACTGCAACCCAAACCTACCATCTTCCCTGAGAACTTAAAGCAACCAATATTG 1261
 DB 1267 CTAGGCTGCAAGTCCCATGATCAACATCTCCAGGAGGAGGAGGAGGAGGAGGAGGAGG 1326
 QY 1262 TGTCTTAAAG-----ACCCCTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1315
 DB 1327 TGGCCCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1386
 QY 1316 CAGAAAG-----ATGTGGAATTTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1354
 DB 1387 CAGAAAGTTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1446
 QY 1355 AAATGTTTCAAGGTTTAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1414
 DB 1447 AAGTGTGCCCCAATGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1506
 QY 1415 CACCCCTAC 1422
 DB 1507 CCACTCC 1514

RESULT 11
 BC041664
 LOCUS
 DEFINITION
 Homo sapiens vitelliform macular dystrophy (Best disease, bestrophin), mRNA (cDNA clone MGC:47884 IMAGE:5194649), complete cds.
 BC041664
 VERSION
 BC041664.1 GI:27371319
 KEYWORDS
 MGC.
 SOURCE
 Homo sapiens
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 2435)
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenman, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., McKusick, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Uedin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Lounellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Wortley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hu, Y., Gibbs, R.A., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fane, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, U., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Small, D.E.,

TITLE
Schnerch, A., Schein, J. E., Jones, S. J., and Marra, M. A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL
MEDLINE
22388257
12477932
2 (bases 1 to 2435)
Straube, R.
Direct Submission
Submitted (20-DEC-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK
COMMENT
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: gcgaps-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nih.gov
Alster, N., Ayle, K., Beckstrom-Sternberg, S. M., Benjamin, B.,
Blakesley, R. W., Bouffard, G. G., Breen, K., Brinkley, C., Brooke, S.,
Dietrich, N. L., Granite, S., Guan, J., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, F., Legaspi, R.,
Maduro, Q. L., Masiello, C., Maekel, B., Mastrian, S. D., McCloskey, J. C.,
McDowell, J., Pearson, R., Stanek, S., Thomas, P. J., Touchman, J. W.,
Taurgeon, C., Vogt, J. L., Walker, M. A., Wecherby, K. D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E. D.

FEATURES
source
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAK Plate: 82 Row: d Column: 4.
Location/Qualifiers
1..2435
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:47884 IMAGE:5194649"
/tissue_type="Brain, adult, 6 pooled whole brains"
/clone_lib="NIH MGC_114"
/lab_host="MDH10B"
/note="Vector: pCMV-SPORT6"
1..2435
/gene="VMD2"
/note="synonyms: BMD, BEST"
/db_xref="LOCUSID:7439"
/db_xref="MIM:153700"
87..1901
/gene="VMD2"
/codon_start=1
/product="VMD2 protein"
/protein_id="AAH41664.1"
/db_xref="GI:27371320"
/db_xref="LOCUSID:7439"
/db_xref="MIM:153700"
/translation="MFEKLTLYCDSYQLIPISFVLGFRVTLVTRWMNOYENLPWPD
RLMSVGFVGEKDEQRLRLTLRYANLGNVLISVSAYVKKRPSAOHLVQAF
MTPAEHOLEKLSLPHNMFWPMVPMFLSKAWLGGIRDPZILLOSILNMTLQAF
CGHLYXWDWISLPLVYQVYVAVYSPFLSLVGRORLPNPKAKGHEHLLVVPFTE
LQPFYVGLKVAEQLINPGEDEDDDEFTNTIVRNLOVSLAVDENHODLPNRPDM
YNNRPPQPTAASQAFRRVSMGSTFNISLKEEMEPQPNQDESDAAGLIGRL
GLQSHDHPRANRRTKLMPKRESLHEGLPKNKAQKQNRQEBNKAAGLAVNA
FKSPLDYQRPYUAPQPLSPMPFLPSPASKLHSYTGIDKDKSLTVSGAK
KSFPLSESDALMEHPEVSQRRKTVFENLDMPEIPEHNLKPELSPFNHTTK
DHMDPVALNRSVLIHNOGCHLACPRPALSILPRLNHLFPHHCQSLDLRLPL
AKGILYLTFTGILKSCSPFLTSPWYHEDFLGREGR"
90..1022
/gene="VMD2"
/note="DUF289; Region: Putative membrane protein. This

family called family 8 in, may be a transmembrane protein
The specific function of this protein is unknown"
/db_xref="CDD:pfam01062"

ORIGIN
Query Match 40.6%; Score 777.2; DB 9; Length 2435;
Best Local Similarity 76.9%; Pred. No. 1e-178;
Matches 1004; Conservative 0; Mismatches 263; Indels 39; Gaps 3;

144 CCATCCGCGACCTCTACAGAAATGGTCTCGAGATGATCAGAGCTGTTGTTGAAGC 203
40 CCCACCTGGAACCCACCTGCTGGCCCTCAGAGAAACACAGCTGATGTTGAGAAC 99
204 TGGCTGTGACTGCGACAGCTACATTCAGCTCATCCCTATATCTTCGTTGGATTCT 263
100 TGACTCTGTATGAGACAGCTACATTCAGCTCATCCCATTTCTCGTGTGCTGGCTTCT 159
264 ATGTTACATTTGGTGTGAGCCGCTGTGTGAGCCAGTACAGAACTTGGCGCCGACC 323
160 ACGTAGCGCTGTGTGACCCGCTGTGTGAAACAGTACAGAACTTGGCGCCGACC 219
324 GCTTCATGATCCAGTGTGTGATGCTTGTGTGAGGGCAAGATGAGGAAGCCGTTTGTGCG 383
220 GCTTCATGAGCTGTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 279
384 GGGCCACGCTCATCCGCTACAGCCATCTCGGCGCAAGTCTCATCTCGCGACATACAGA 443
280 GGGCCACGCTCATCCGCTACAGCCATCTCGGCGCAAGTCTCATCTCGCGACGCTACAGA 339
444 CTTGCGTCTACAGCGCTTCTCCACTTTCACCACTGTGTGTACAGAGTTTATGACCC 503
340 CCGAGCTTACAGCGCTTCTCCAGCGCCGACGACCTGTGTGCAAGCAGGCTTATGACTC 399
504 ATGGGAAACATTAACAGTGTGAGAGTGGGCTTACCAACAACAACATTCGGGCGCT 563
400 CGGAGAAACATTAACAGTGTGAGAGTGGGCTTACCAACAACAACATTCGGGCGCT 459
564 GGGTGTGTTTGGCACTTGTCAATGAAAGGCTTATCTTGGAGTGTGAATCCGGACACCG 623
460 GGGTGTGTTTGGCACTTGTCAATGAAAGGCTTATCTTGGAGTGTGAATCCGGACACCTTA 519
624 TCTGTCTCAGAGCTTGAATGAGTGTGTACTTTCGTAATCTGAGTGTGACAGCTGT 683
520 TCTGTCTCAGAGCTTGAATGAGTGTGTACTTTCGTAATCTGAGTGTGACAGCTGT 579
684 ATGCTTACAGTGTGATTAAGTATCCCATTTGGTGTACACAGTGTGTGACATGTGACAGTAT 743
580 ATGCTTACAGTGTGATTAAGTATCCCATTTGGTGTGTGACAGTGTGTGACATGTGACAGTAT 639
744 ACAGCTTTTCTTCTGCAATGCTGTGATCGGAGGAGGAGTGTGACACCAACAAGAGCTACG 803
640 ACAGCTTTTCTTCTGCAATGCTGTGATCGGAGGAGGAGTGTGACACCAACAAGAGCTACG 699
804 CAGGCAATGAGATGATGTGTTGTGCTGTCTTCAACATCTGCAATCTTATTTATCA 863
700 CTGGCAATGAGATGATGTGTTGTGCTGTCTTCAACATCTGCAATCTTATTTATCA 759
864 TGGCTGTGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 923
760 TGGCTGTGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 819
924 TTTGAGCAATGAGTGTGATGATTTGAGCAAGAACTGAGAGTGTGAGTGTGAGTGTGAGTGT 983
820 TTTGAGCAATGAGTGTGATGATTTGAGCAAGAACTGAGAGTGTGAGTGTGAGTGTGAGTGT 879
984 TGCACCAAGATTTGCTTCCATGAGAACTGATCATGTATGAGAACTGAGAACTGAGAACT 1043
880 TGCACCAAGATTTGCTTCCATGAGAACTGATCATGTATGAGAACTGAGAACTGAGAACT 939
1044 CGCCTTACAGCTGTGCTTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1103
940 CCCCTTACAGCTGTGCTTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 999

QY	1104	ACATCAGCCTAAAGAAAGAAAGCTTAGAGCTTTGGTCCAAAGAGAGAGGCTACACGGATA	1163
Db	1000	ACATCAGCCTAAAGAAAGAGGAGTAGGAGTTCCAGCCCAATCCAGAGAG-----	1047
QY	1164	AGAAAGAGTAGGCGCTATAGACAGCACATTAGGCTGGCTTTTGGATGTGACACCCAAAAC	1227
Db	1048	ACGAGGAAGATGCTCAACGCTGGCATCATATTGGCCGGCTCTTAGCGCTGACATGCCATGATC	1107
QY	1224	ACCATCTTCCCTTGGAAAGACTTAAAGACCAAACTATTGTGTTCTTAAAG-----ACCCCC	1277
Db	1108	ACCATCTCTCCAGGGCAAACTCAGAGACCAAACTACTGTGGCCCAAGAGGAATCCCTTC	1167
QY	1278	TCCTTGAAGGCCAGTGTAGGATGCGCAACAGAAAAACAGAAAG-----	1322
Db	1168	TCCACGAGAGGCGTGGCCAAAAACCAAGAGGACGCAAAACAGATGTTAGGGGCCAGAG	1227
QY	1333	-----ATGCTCGAAATTTAAGGCTCGGACTTTGAAAATGTGTTCCAAAGTTTAAAG	1376
Db	1228	ACAAACAAAGGCGTGGAAAGCTTAAAGGCTGTGGACGCTTCAAGTCTGCCCACTGATTCAGA	1287
QY	1377	GGAAGAGGTCCCATTTGTGGCCCAACAGGACCCAGAGACCCACTTAC	1422
Db	1288	GGCCAGGCTACTTACAGTGCCTCCACAGAGGCCCTCTAGCCCACTCC	1333

	RESULT 12		
LOCUS	Af057170	2420 bp	mRNA linear PRI 17-OCT-1998
DEFINITION	Homo sapiens bestrophin (VMD2) mRNA, alternatively spliced product,		
ACCESSION	Af057170		
VERSION	Af057170		
KEYWORDS	GI:3335160		
SOURCE			
ORGANISM	Homo sapiens (human)		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	Petrushkin,K., Koist,M.J., Bakall,B., Li,W., Xie,G., Martnell,T., Sandgren,O., Forsman,K., Holmgren,G., Andraasson,S., Vujic,M., Bergen,A.A., McGarty-Dugan,V., Figueroa,D., Austin,C.P., Metzker,M.L., Caskey,C.T. and Madellis,C. Identification of the gene responsible for Best macular dystrophy Nat. Genet. 19 (3), 241-247 (1998) 98324772		
MEDLINE	98324772		
PUBMED	9662395		
REFERENCE	2 (bases 1 to 2420)		
AUTHORS	Petrushkin,K.		
TITLE	Direct Submission		
JOURNAL	Submitted (03-APR-1998) Human Genetics, Merck Research Laboratories, West Point, PA 19486, USA		
FEATURES			
source	location/Qualifiers		
	1..2420		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/chromosome="11"		
	/map="11q13"		
	1..2420		
	/gene="VMD2"		
	105..1412		
	/gene="VMD2"		
	/note="Best's macular dystrophy gene; BMD; alternatively splice isoform; contains exon 7"		
	/codon_start=1		
	/product="bestrophin"		
	/protein_id="AAC64344.1"		
	/db_xref="GI:3335161"		
	/translation="MTITTVSYOVNARLGSFRLLCWRGSIYKLVGEPLIFLCYY IIPFLATLTBEQQLMEFKTLICDSYILDLPISFYIGPYTLLVTRMNNOVENLPM PDLMSIVSGFVEKDRQGSLRLRTLIIRANLGVLILRSVSTRVYKRFSAGHLVOA GPPTPAHKOLEKLSLEPHNNFWPWFAWLMKAWLGGISRTDPILOSILNENNTLR TCGCHLAYOWMISIPLYVTGVTVVAVSFFLTCAVGQFLNPAAKGHELDLVVPVF		

ORIGIN

TELQFFFYVWGLKVGSLRALGWRHGORHGOQLLETROMCOERKYSRVRESSOAMWRRT
PVIPTAREAGESLEPGRRLWMOSSSTPLERMMILRPTGLSTGICRCPCLMLMR
CTRCLGWSRTCTGISPSHSPPTQLLPSSVEPPLMAPPSTA "

Query Match	34.4%	Score 658.2;	DB 9;	Length 2420;
Best Local Similarity	67.6%	Pred. No. 1.3e-149;		
Matches 116;	Conservative 0;	Mismatches 293;	Indels 242;	Gaps 4

QY	2	TGCCAAGCATGACTATATACCTACACAAACAAAGTAGCCAAATGCCCGCTCGGTTTCGTTT	61
Db	96	TGCCGTGGCAATGACATCATCTTACACAAGCCAAAGTGGCTAATATGCCGCTTAGGCTTC	155
QY	62	TGTCGCCCTCTCTGTCGTGGCGAGGCGACATCTACAAGCTGCTGTATAGAGAAATTCCTT	121
Db	156	TCCCGCCTGCTGCTGTGTGTGGCGGGGAGCATCTACAGCTGTCTATATGGCGAATTCCTTA	215
QY	122	GTCTTCATATTTCTCTACTATATTCATCCGTGGAACTCTACAGAAATGGTTCTCTGAGTAT	181
Db	216	ATCTTCCTGCTCTGCTACTACTATCATTCGCTTTATTTATTAAGCTGGCCCTCACGGAGAA	275
QY	182	CAGCAGCTGTGTGTTGAGAAAGTGGCTCTGTACTACGACAGCTACTATCAGTCAATCCCT	241
Db	276	CAAAAGCTGATGTTTGAAGAACTAATCTGTATTTGCGAACACTATCATCAGCTCATTCCTC	335
QY	242	ATATCTCTCGTCTGGGTTTCTATGTATACATTTGGTGGAGCCGCTGGTGGAGCCAGTAC	301
Db	336	ATTTCCTTCGTGCTGGGCTTTCTAAGTACGCTGGTGTGACCCGCTGGTGAACCAATAC	395
QY	302	GAGAACTTGCCGTGGCCGACCGGCTCATATGATCAGAGTGTAGCTTGTGAGAGGGCAAG	361
Db	396	GAGAACTTGCCGTGGCCGACCGGCTCATATGATCAGAGTGTAGCTTGTGAGAGGGCAAG	455
QY	362	GATAGAGAAAGGCGTTTGTCTGGCGGCGACAGCTCATCCGCTACCGCATCTGGGCCCAATG	421
Db	456	GACAGAGAAAGGCGGCTGTGGCGGCGACAGCTCATCCGCTACCGCATCTGGGCCCAATG	515
QY	422	CTCATCTGCGACATACAGCACTCGGCTGTACAAAGCGCTTCCCACTCTTCAACCACTG	481
Db	516	CTCATCTGCGACAGCTACAGACCGGCACTCTACAAAGCGCTTCCCAAGGCCCAAGACTG	575
QY	482	GTGCTAGCAGGTTTATATGACCCATGGGAAACATTAAGCAGTTGCAGAAATGGGCTTACCA	541
Db	576	GTGCAACAGGCTTTATATGATCTCCGGCAGAAACAAAGCAGTTGGAAGAACTAGACTTACA	635
QY	542	CACAAACATTTTGGGTGCTCGGCTGTGGTTTGCCAACTTGCTATTAAGACCTATATT	601
Db	636	CACAAACATTTTGGGTGCTCGGCTGTGGTTTGCCAACTTGCTATTAAGACCGCTGTG	695
QY	602	GGAGGTGGAATCCGGGACACCGTCTGTCTCCAGAGCCGTAGTAATGAGGTGTGTACTTGG	661
Db	696	GGAGGTGGAATCCGGGACCGCTATCTCTGCTCCAGAGCTGTGCTGAACGAGATTAACACTTG	755
QY	662	CGTACTCAGTGTGACAGCTGTATGCTCTACGACTGATTAAGTATCCATTGGTGTACACA	721
Db	756	CGTACTCAGTGTGACAGCTGTATGCTCTACGAGCTGATTAAGTATCCACTGTGTATACCA	815
QY	722	CAGGTGTGACAGTGGCAGTATATACGCTTTTCTTGATAGCTTGTAATCGGAGGCAATTT	781
Db	816	CAGGTGTGACAGTGGCAGTATATACGCTTTTCTTGATAGCTTGTAATCGGAGGCAATTT	875
QY	782	CTGAACCCAAACAAGAGTACCCAGGGCATAGATAGATCTGTTGGGCTGTCTTACACA	841
Db	876	CTGAACCCAAACAAGGCTTACCCCTGGGCAATAGAGTGGACCTGTGTGCTTACACG	935
QY	842	ATCCGCAATTTCTATTCTACATAGGAGCTGACTGAA	876
Db	936	TTCCTGTAGATTCTTCTTCTATATGTTGGCTGAGTGAAGTGGGCTCTCAAGGACCCTGTG	995
QY	877	-----	876
Db	996	GGCTGAGAGCATGGCCAGAGGGGTCAATGGCCAGCAGCTGTGTGAGAGAGATCAGTGT	1055

```
QY 877 ----- 876
Db 1056 CAGGAAAGAGGTCTCA CGGGTAGAAAGACAGCGCGTGTGCGCACACTGTATC 1115
QY 877 ----- 878
Db 1116 CCAGCTACTCGGAGGCTGAGGACAGAGATCGCTTGAACCCGGAGCGGAGGTTGTGG 1175
QY 879 TGGCAGAACGCTCATCAACCCCTTCGGGAGAGACGATGATGATTTTGAAGTAACCTGA 938
Db 1176 TGGCAGAGAGCTCATCAACCCCTTCGGGAGAGATGATGATTTTGAAGCAACTGGA 1235
QY 939 TCATTGACAGAAACCTCGAGGTGTCCCTGTTCCTGTGATGAGATGACCAAGACTTGC 998
Db 1236 TTTCGACAGAGATTTTGACAGGTGTCCCTGTTCCTGTGATGAGATGACCAAGACTGC 1295
QY 999 CTCCTCATGGAACGTGACATGTACTGGAACGAGGAGCGCTTCAGCCGCTTACACAGTG 1058
Db 1296 CTCGGATGAGACCGGACATGTACTGGAATTAAGCCGAGCCACAGCCCTTACACAGCTG 1355
QY 1059 CTTCTGCGAGGTCTCGCCGGGATTCCTTCATGGGCTCCACCTTCAACATCAGCCTAAAGA 1118
Db 1356 CTTCCGCCCACTTCCTGTGAGCTTCCTTTATGGCTCCACCTTCAACATCAGCCTGAACA 1415
QY 1119 AAGAAAGCTTGAAGCTTGTGTCAAAAGAGAGCTGACACGAGATTAAGAAAGAGTGTCT 1178
Db 1416 AAGAGAGATGAGATTCAGCCCAATGAGAG-----AGAGAGAGATGTCTC 1463
QY 1179 ATGACAGACCATATGAGCTGCTTTTGAAGCTGCAACCAAAACTACCATTTCCCTTGA 1238
Db 1464 ACGCTGCATCATTTGGCGCTTCCTAGGCTGACAGTCCCATGATCACCATCTCCCAAGG 1523
QY 1239 AAGACTTAAGACCAAACTATTTGTCTTAGAA-----CCGCTCCTCGAGAGCCAGT 1292
Db 1524 CAAACTCAAGACCAAACTACTGTGCGCCCAAGAGGAAATCCCTTCTCCACAGAGGCTGC 1583
QY 1293 GTAAGATGCCAACAGAAAAACAGAAAGAT-----GCTTGA 1331
Db 1584 CCAAAACCAAGAGCAGCCAAACAGAAAGTTAGGGGCAAGAAAGAACAAAGGCTTGA 1643
QY 1332 AATTAAAGGCTGTGACTTTTGAATGTGTTCCAGGTTTAAAGAGAGGCTTCCATT 1391
Db 1644 AGCTTAAGGCTGTGAGCGCTTCAAGTCTGCCCACTGATCAAGAGGCAAGGCTACTACA 1703
QY 1392 GTGGCCACAGGCAAGCCACAGGCAACCTTAC 1422
Db 1704 GTGCCACACAGCGCCCTCAGCCCACTCC 1734
```

```
RESULT 13
BD136721
LOCUS BD136721 2429 bp DNA linear PAT 18-SEP-2002
DEFINITION Beak's macular dystrophy gene.
ACCESSION BD136721
VERSION BD136721.1 GI:23231666
KEYWORDS JP 2002504559-A/3.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2429)
Petrushkin, K., Gabkey, T.C., Metzker, M. and Madeline, C.
Beak's macular dystrophy gene
Patent: JP 2002504559-A 3 12-FEB-2002;
MERCK & CO INC, CLAS WADLIUS
OS Homo sapiens (human)
PN JP 2002504559-A/3
PR 12-FEB-2002
PR 25-FEB-1998 JP 2000533447
PR 25-FEB-1998 US 60/075941, 18-DEC-1998 US 60/112926 PI
KONSTANTIN PETRUSHKIN, THOMAS C GABKEY, MICHAEL METZKER, CLAES PI
WADLIUS
COMMENT
JOURNAL
```

```
REFERENCE
AUTHORS
TITLE
JOURNAL
PC C07K16/18, C07K14/47, C12N5/10, C12N15/09, C12P19/34, C12Q1/68// PC
```

```
C12P21/08,
PC C12N5/00, C12N15/00
CC Beak's macular dystrophy gene
FH Location/Qualifiers
FT source 1..2429
FT 1..2429 /organism='Homo sapiens (human)'
FEATURES
source
1..2429
/organism='Homo sapiens'
/mol_type='genomic DNA'
/db_xref='taxon:9606'
ORIGIN
Query Match 34.4%; Score 658.2; DB 6; Length 2429;
Best Local Similarity 67.6%; Pred. No. 1,36-149;
Matches 116; Conservative 0; Mismatches 293; Indels 242; Gaps 4;
```

```
QY 2 TGGCAAGCATGACTATACCTTACCAACAAAGTACCAATGCCGCTCGTTCCGTTTC 61
Db 96 TGCTGCTCATGACATCACTTACCAAGCCAAAGTGTGCTAATGCCGCTTACGCTTC 155
QY 62 TCGTCCCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 121
Db 156 TCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 215
QY 122 GCTTCATATTCCTCTACTATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 181
Db 216 ATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 275
QY 182 CAGCAGCTGTTGTTTGAAGAGCTGCTGTACTGCGACAGCTACATTCAGCTCATCCCT 241
Db 276 CAACAGCTGATGTTTGAAGAACTGACTGTGATGCGACAGCTACATTCAGCTCATCCCT 335
QY 242 ATATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 301
Db 336 ATTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 395
QY 302 GAGAACTTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 361
Db 396 GAGAACTTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 455
QY 362 GATGAGAAAGGCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 421
Db 456 GACGAGCAAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 515
QY 422 CTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 481
Db 516 CTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 575
QY 482 GTGCTAGAGGTTTATGACCCATGAGGAAACATAGAGAGTTGCAAGTTGGGCTTACA 541
Db 576 GTGCTAGAGGTTTATGACCCATGAGGAAACATAGAGAGTTGCAAGTTGGGCTTACA 635
QY 542 CACAACATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 601
Db 636 CACAACATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 695
QY 602 GAGAGTGAATCCGGACACCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 661
Db 696 GAGAGTGAATCCGGACACCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 755
QY 662 GTATCTGAGTGGACAGCTGATAGCTTACGATGAGTAAAGTAAAGTAAAGTAAAGTAA 721
Db 756 GTATCTGAGTGGACAGCTGATAGCTTACGATGAGTAAAGTAAAGTAAAGTAAAGTAA 815
QY 722 CAGGTGTGACAGTGGACATACAGCTTTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 781
Db 816 CAGGTGTGACAGTGGACATACAGCTTTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 875
QY 782 CTGAACCAAAACAGAGCTACCAAGGCAATGAGTGAATCTGCTGCTGCTGCTGCTGCT 841
Db 876 CTGAACCAAAACAGAGCTACCAAGGCAATGAGTGAATCTGCTGCTGCTGCTGCTGCT 935
```

Db	842	ATCTGCAGATTTCTATTCTTATCATGAGGCTGGCTGAA-----	876
Qy	842	ATCTGCAGATTTCTATTCTTATCATGAGGCTGGCTGAA-----	876
Db	936	TTCTCTGACGATTTCTTCTTATGTTGGCTGGCTGAAGTGGCCCTCTCCAGGAGCCCTGCTG	995
Qy	877	-----	876
Db	996	GGCTGAGAGCATGGCCAGAGGGGATCATGGCCAGCAGCTGTTGAGACGAGGATGACGATGT	1055
Qy	877	-----	876
Db	1056	CAGAAAAGAAAGTCTCACGGGTAGAAAAGCAGCCAGCGGTGATGGCGACACTGTATC	1115
Qy	877	-----GG 878	878
Db	1116	CCAGCTACTCGGAGAGGCTGAGGCGAGAGAAATCGCTTGAACCCGGAGGCGGAGGTTGTGG	1175
Qy	879	TGGCAGAAACAGCTCATCAACCCTTCGGGAGAGACGATGATGATTTTGAGCTAATTGA	938
Db	1176	TGGCAGAGCAGCTCATCAACCCTTTGGAGAGGATGATGATGATTTTGAGCAACAATCGA	1235
Qy	939	TCATTGACAGAAACCTGCAGGTTTCCTGTTGTCCGTGGATGGATGACCAAGAACTTGC	998
Db	1236	TTGTGACAGAAATTTGGCAGGTGTCCCTGTTGGCTGTGTGAATGAGATGACCAAGGACCTGC	1295
Qy	999	CTCCCATGGAAAGTGAACAATGTAAGTGAACGAGGCGACGCGCTCAGCGCGCTTACACAGCTG	1058
Db	1296	CTCGAATGAGACCGGACATGTACTGGAATTAAGCCGAGCCACAGCCCTTACACAGCTG	1355
Qy	1059	CTTCTGCAGAGTCTGCGCGGATTCCTTCATGGGCTCCACCTTCMAATCAGCTTAAGA	1118
Db	1356	CTTCCGCCCAATTCGCTGAGCCCTCTTTATGGGCTCCACCTTCMAATCAGCTTAAGA	1415
Qy	1119	AAGAAAGCTTAGAGCTTTGTGCAAAAGAGAGCTGACACGGATTAAGAAAGAGTGGCT	1178
Db	1416	AAGAGGAGATGAGATTCAGGCCCAATCAGGAGG-----ACAGAGAGATGTCTC	1463
Qy	1179	ATTACGAGACCATAGGCTGCTCTTGAAGACTGCAACCCAAACATACATCTTCCCTTGA	1238
Db	1464	ACGCTGGATCATTTGGCCGCTTCTTAGGCTGACGTCCCATGATCATCATCTTCCAGGG	1523
Qy	1239	AAGACTTAAAGACCAAACTATTTGTTGTTTAAGAA-----CCCTCTCCTGAAAGCCAGT	1292
Db	1524	CAAACTCAGAGACCAAACTACTGTGTGCCCAAGAGGAGATCCCTTCTCAGCAGAGGCTCG	1583
Qy	1293	GTAAGATGCGAACCCAGAAAAACCAAGAAAGAT-----GTCTGGA	1331
Db	1584	CCAAAAACCAAGAGCAGCCAAACAGACGTTAAGGGGCGAAGAAAGACAACAAAGGCTTGA	1643
Qy	1332	AATTTAAGGCTCTGACATCTTTGAATATGTGTCCAAAGTTTAAAGAGAGAGGCTCCCAT	1391
Db	1644	AGCTTAAAGGCTGTGAGCGCTTCAAGTGTGGCCCATGTATTCAGAGGCGCAGGCTACTACA	1703
Qy	1392	GTGCGCCACAGAGCACCCGAGCCACCTTAC	1422
Db	1704	GTGCGCCACAGAGCGCCCTCAGGCCCACTCC	1734
RESULT 14			
BC015220			
LOCUS			
DEFINITION			
BC015220 2441 bp mRNA linear PRI 07-OCT-2003			
Homo sapiens vitelliform macular dystrophy (Best disease, bestdrophin), mRNA (CDNA clone MGC:17761 IMAGE:3877806), complete cds.			
ACCESSION			
BC015220			
VERSION			
BC015220.2 GI:34783515			
KEYWORDS			
MGC.			
SOURCE			
Homo sapiens (human)			
ORGANISM			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE			
1 (bases 1 to 2441)			
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, J., Shenmen, C.M., Schueller, G.D., ..,			

JOURNAL	TITLE
MEDLINE	
PUBMED	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
REMARK	
COMMENT	

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
On Sep 16, 2003 this sequence version replaced gi:21955361.
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: DCTD/DTP/Gazdar
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www.sngc.stanford.edu>
Contact: (Dickson, Mark) mcd@pxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAX Plate: 14 Row: b Column: 9.

FEATURES	SOURCE
source	
gene	
CDS	

```

1 . .2441
   /organism="Homo sapiens"
   /mol_type="mRNA"
   /db_xref="taxon:9606"
   /clone="MGC:17761 IMAGE:3877896"
   /tissue_type="lung, large cell carcinoma"
   /clone_id="NIH_MGC_68"
   /lab_host="DH10B"
   /note="Vector: pCMV-SPORT6"

1 . .2441
   /gene="VMD2"
   /note="synonyms: BMD, BESt"
   /db_xref="LOCusID:7439"
   /db_xref="MIM:153700"
   573 ..2069
   /gene="VMD2"
   /codon_start=1
   /product="VMD2 protein"
   /protein_id="AAH15220.1"
   /db_xref="GI:22800661"
   /db_xref="LOCusID:7439"
   /db_xref="MIM:153700"
   /translation="MEKTLRLCYDSYITQLIPISFVLGFVTYLVTTRKNQYNLPDD
RMSLSGVPEKDQEGRLRLRTLIRANIGNVALRSVSAYAKRPSSAOHLVDQAFG
MTPAEHKOELSLPHNMFWPMWNEISMKMLGRIDPILLOSLLBMENLRQO
CGHLYAVDWISTPLVTOVVTVAVSFCLCGLRGOLINPAKAVPGHELDLVVFTEF
LOFPFYGNLKSVLAVDENHQDLPRNEPDMYNNKPQPPTYAASQFRASMGST

```


FNISLNKEMEFQNOEDBEDAHAGIIGRFLQSHDHPHPRANSRTKLMPKRESIL
HEGLPKHKAQKONVGOEDBNKAWLKAVDPFKAPYQRYGYSAQOTPLSPMPFE
PLEPSAPKLSHVSVDITKDKSLKTVSSGAKSFELISESGALMEHPVSQVARKTV
ERNITDMEIPENHLKEPLESPTNIHTLKDHPDYMALENRDEAS"
576..1427
/gene="VMD2"
/note="DUF289; Region: Putative membrane protein. This
family called family 8 in, may be a transmembrane protein
The specific function of this protein is unknown"
/db_xref="CDD:pfam01062"

ORIGIN

```
Query Match      32.6%; Score 624.4; DB 9; Length 2441;
Best Local Similarity 71.6%; Pred. No. 2.3e-141;
Matches 935; Conservative 0; Mismatches 251; Indels 120; Gaps 4;
```

144 CCATCCGTGAGCTTACAGAAATGGTCTCTCGAGTATCAGCAGCTGTTGTTAGAAAGC 203
Db CCCACCTGGAACCCACCTGCTGCCCCCTCAGGAAAGAACACACTGATGTTGAGAAAC 585
Qy TGGCTCTGTAAGGAGAGCTTACATTCATCCCTATCCCTTCTGCTGCTGCTTCT 263
Db TGAATCTGTAATGAGACAGCTACATCAGCTATCCCATTTCTTGTGTGGGCTTCT 645
Qy ATGTTAATTTGGTGTGAGCCGCTGTGTGAGACAGTACAGAACTTGGCCGTGGCCGAC 323
Db ACGTGAAGCTGTGTGTAACCCGCTGTGTGAGACAGTACAGAACTTGGCCGTGGCCGAC 705
Qy GCCTCATGATCAGAGTGTCTAGCTTCTGTGAGGGCAAGATGAGAAAGCCGTTTGTCTG 383
Db GCCTCATGAGCTGTGTGTGCGGCTTCTGTGAGGGCAAGATGAGAAAGCCGCTGTCTG 765
Qy GGGCCAGCTCATCCGCTACAGCCATCCGAGGCAAGTCTCATCTGGGAGCATTCAGCA 443
Db GGGCCAGCTCATCCGCTACAGCCATCCGAGGCAAGTCTCATCTGGGAGCATTCAGCA 825
Qy CCTGCTCTACAAAGCGCTTCCCACTCTTACCAAGCTGTGTGAGAGTCTTATGACCC 503
Db CCGGAGCTTACAAAGCGCTTCCCAAGCGCCAGCACTGTGTGAGAGCTTATGATC 885
Qy ATGAGGAAATATAGCAGTTGAGAAAGTTGGGCTTACCAACAGCAATCTTGGGTGCCCT 563
Db CGGAGAAACATAGCAGTTGAGAAAGTTGGGCTTACCAACAGCAATCTTGGGTGCCCT 945
Qy GGGTGTGTGTTGGCCAACTGTGCAATGAAAGGCTTGAAGGTCGAAATCCGGGACCTTA 1005
Db GGGTGTGTGTTGGCCAACTGTGCAATGAAAGGCTTGAAGGTCGAAATCCGGGACCTTA 1065
Qy TCCTGCTCCAGAGCCTGATGAAATGAGGTTGTAACCAAGTGTGACAGTGT 683
Db TCCTGCTCCAGAGCCTGATGAAATGAGGTTGTAACCAAGTGTGACAGTGT 1065
Qy ATGCTTACGACTGATATGATATCCATTTGATGTAACACAGTGTGACAGTGT 743
Db ATGCTTACGACTGATATGATATCCATTTGATGTAACACAGTGTGACAGTGT 1125
Qy ACAGCTTTTCTTCTGCACTGTTGATGAGGAGGAGTTTCTGAACCAACAGAGCTATCC 803
Db ACAGCTTTTCTTCTGCACTGTTGATGAGGAGGAGTTTCTGAACCAACAGAGCTATCC 1185
Qy CAGGCAATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 863
Db CAGGCAATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1245
Qy TGGGCTGAGTGAAGTGGCAGAACAGCTCATCAACCCCTTGGGAGAGAGATGATGAT 923
Db TGGGCTGAGTGAAGTGGCAGAACAGCTCATCAACCCCTTGGGAGAGAGATGATGAT 1257
Qy TTGAGACTAATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 983
Db TTGAGACTAATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1284
Qy TGACACGAACTTGCTCCATGAGAAAGTGAATGATGATGATGATGATGATGATGATGAT 1043

1285 TGACACGAGACCTGCTCCGATGAGACCGGACATGATGAAATAGACCCGACAGC 1344
Qy GGCCTTACACAGCTGCTTCTGCGCAGCTGTGCGCCGACATCTCTTATGAGGCTTCA 1103
Db CCCCTTACACAGCTGCTTCTGCGCAGCTGTGCGCCGACATCTCTTATGAGGCTTCA 1404
Qy ACATCAGCTTAAAGAAAGACTTAGAGCTTGTGTCAAAAGAGAGCTGACACGATA 1163
Db ACATCAGCTTAAAGAAAGACTTAGAGCTTGTGTCAAAAGAGAGCTGACACGATA 1452
Qy AGAAAGAGAGTGTATAGCAGACATAGGCTGCTTCTTGAAGCTGACCAAAACT 1223
Db ACGAGAGAGATGCTACAGCTGATCATTTGGCCGCTTCTTGAAGCTGACATGATC 1512
Qy ACCATCTTCCCTTAAAGACTTAAAGCAAACTATGTTGTTCAAA-----ACCCCC 1277
Db ACCATCTTCCCTTAAAGACTTAAAGCAAACTATGTTGTTCAAA-----ACCCCC 1572
Qy TCTCGAAGGCAATGTAAGATGCAACAGAAACAGAAAG----- 1322
Db TCTCGAAGGCAATGTAAGATGCAACAGAAACAGAAAG----- 1632
Qy -----ATGCTGGAATTTAAGGCTGGAATCTTGAATGTTTCAAGTTTAA 1376
Db ACAACAGGCTGGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAG 1692
Qy GGAAGGCTCCCATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1422
Db GGCAGGCTTACTACAGTGCCTCCACAGACACCTTCAAGCCCATCTC 1738

RESULT 15
AX745966 2137 bp DNA linear PAT 14-MAY-2003
LOCUS AX745966
DEFINITION Sequence 3 from Patent WO03030922.
ACCESSION AX745966
VERSION AX745966.1 GI:30724620
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1
AUTHORS Steuernagel, A., Bruemer, G., Fritsch, R., Eulenber, K. and
Closser, T.
TITLE Beetrophin and beetrophin homologous proteins involved in the
regulation of energy homeostasis
JOURNAL Patent: WO 03030922-A 3 17-APR-2003;
Developen Aktiengesellschaft fuer Entwicklungsbioologische Forschung
(DE)
FEATURES
source location/Qualifiers
1..2137
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

```
Query Match      27.1%; Score 519.4; DB 6; Length 2137;
Best Local Similarity 66.9%; Pred. No. 1.1e-115;
Matches 755; Conservative 0; Mismatches 371; Indels 3; Gaps 1;
```

9 CCATGACTATCACTTACAAACAAAGTAGCAATGCCGCTGCTGCTGCTGCTGCTGCTGCT 68
Db CCATGACTATCACTTACAAACAAAGTAGCAATGCCGCTGCTGCTGCTGCTGCTGCTGCT 269
Qy TCCCTCTGTGCTGGCGAGGCGAGCTTACAAAGCTGCTGTAAGAGAAATCTTGTCTTCA 128
Db TCCCTCTGTGCTGGCGAGGCGAGCTTACAAAGCTGCTGTAAGAGAAATCTTGTCTTCA 329
Qy TATTCCTTACTATTTCCATCCGTGAGCTTACAGAAATGTTCTTCTGAGTATGAGCAGC 188
Db TATTCCTTACTATTTCCATCCGTGAGCTTACAGAAATGTTCTTCTGAGTATGAGCAGC 389


```
QY 189 TGTGTTTGAGAGCTGGCTCTGTACTGCGACAGCTACATTCACTGATCCCTATATCCT 248
Db 390 GCTACTTCGAGAACCTTGATATTTATGTGACAGATATGCCAGCTCATCCCTGTCTCT 449
QY 249 TCGTTCTGGGTTTCTATGTATACATTGGTGGAGCCGCTGGTGGAGCCAGTACGAGAACT 308
Db 450 TCGTGTCTGGCTTTTATGTGACGCTGGTGGTGAACCGCTGGTGGAGCCAGTACCTATGCA 509
QY 309 TGGCGTGGCCGACCGGCTCATGATCCAGGTGTCTAGCTTGTGTGAGAGGCAAGATGAGG 368
Db 510 TGGCGTGGCCGACCGGCTCATGATGCGGTGGGCGGCGACCGTGTGACAGAGCGACGACC 569
QY 369 AAGGCGCTTGTGTGGGCGGACGCTCATCCGCTACGCGCATCCGCGGCAAGTGTCTATCC 428
Db 570 GCGGCGCGCTCTACCGGCGCACTCATGCGCTAGCGAGGAGCTCTCGGCGGTGTCTATCC 629
QY 429 TGGCGAGCATACAGCACTCGGTCTACAGCGCTTTCCCACTCTTCAACCACCTGGTGTAG 488
Db 630 TGGGCTCCGTCAGACACCGCGGTGTCAAGGCTTCCCAACATAGACCAAGTGGTGGAGG 689
QY 489 CAGGTTTTATGACCCATGGGAGACATAGCAGTTGCAAGTTGGGCTTACCACACACA 548
Db 690 CTGGGTTTATGACCCGAGAGAGCGCAAGATTGAAAACCTGAACCTCATCTTACACA 749
QY 549 CATTCGTGGTGGCCCTGGGTGTGTTTGGCACTGTCAATGAGGCGCTATCTTGAAGTTC 608
Db 750 AGTACTGGGTGCCCTGGGTGTGTTTCTCAACTGGCGGCAAGGCGCGAGGAGGCC 809
QY 609 GAATCCGGGACACCGCTGCTCCAGAGCCGTGATGAATGAGGTGTACTTGGGTACTC 668
Db 810 GCATCCCGGACACAGGCGCCCTTAAGCTGTGCTCGAGAGCTGATGTGTTTTCGGGACA 869
QY 669 AGTGTGACAGCTGTATGCTTACAGCTGGATATCCATTTGGTGTACACAGGTGG 728
Db 870 AATGTGGAATGCTCTTTCACATATGATGATTAAGTAACTCCCTGTGTACAGCAAGGTGG 929
QY 729 TGAAGTGGGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 788
Db 930 TGACCATGCACTGTACAGTACTTCTGCGCTTGTGCTCATTTGTGCGCAAGTTCCTGAGAC 989
QY 789 CAAACAAAGACTACCCAGGCGCATGATGATGATGATGATGATGATGATGATGATGATGATG 848
Db 990 CGGCTCAGGGTTACAAAGACCAAGCACTAGACTGTGTGTGCCATCTTCACTCTTGC 1049
QY 849 AATTCTTATTTCAATGGGCTGGCTGAGAGGTGGAGAAAGACTCATCAACCCCTTGGGG 908
Db 1050 AGTCTTCTCTACCGCGGCTGGCTCAAGGTGAGCTGAGCAAGCTCATCAACCCCTTGGAG 1109
QY 909 AGGACGATGATGATTTTGAAGTAACTGATCATGACAGAAACCTGCAAGGTGTCCCTGT 968
Db 1110 AGGACGATGATGATTTTGAAGTAACTTCTGATGATGATGATGATGATGATGATGATGATG 1169
QY 969 TGTCCGTGATGGATGACACAGAACTTGCTCCCATGAAAGTGAATGATGATGATGATGATG 1028
Db 1170 TGGCAGTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1229
QY 1029 AAGCAGCGCTCAGCGCGCTTACAGAGTGTCTG---CCAGGTCTGGCGGCAATTCCT 1085
Db 1230 CAGCCGAGGCTCGCGCCCATACAGAGGCTACGTCTTCCAGCTGCGGAGGCTTCTCT 1289
QY 1086 TCATGGGCTCAGCTTCAACATCAGCTTAAAGAAAGAGACTTGAAGCT 1134
Db 1290 TCCAGGGCTCCACTTTGACATCAGCTGCGCAAGAAAGACATGCAATT 1338
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using bw model

Run on: March 25, 2005, 22:43:02 ; Search time 6161.57 Seconds
(without alignments) 11836.456 Million cell updates/sec

Title: US-09-622-964a-28

Perfect score: 1916
Sequence: 1 gtgcacagcatgactatca.....aaaaaaaaaaaaaaaa 1916

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: gb_esc1:*
2: gb_esc2:*
3: gb_hic:*
4: gb_esc3:*
5: gb_esc4:*
6: gb_esc5:*
7: gb_esc6:*
8: gb_gsa1:*
9: gb_gsa2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1684.8	87.9	1699	3	AK006549	AK006549 Mus muscu
2	573.6	29.9	963	5	B0879880	B0879880 AGENCOURT
3	519	27.1	713	6	BY114331	BY114331 BY114331
4	502.6	26.2	881	6	CD518675	CD518675 AGENCOURT
5	493.6	25.8	792	7	CK470840	CK470840 AGENCOURT
6	488.2	25.5	756	7	CO396179	CO396179 AGENCOURT
7	487	25.4	664	6	BY114344	BY114344 BY114344
8	474.2	24.7	862	7	CK469826	CK469826 AGENCOURT
9	466	24.3	578	1	AU040308	AU040308 AU040308
10	464	24.2	570	1	AU040310	AU040310 AU040310
11	428.2	22.3	844	7	CK653343	CK653343 AGENCOURT
12	397.4	20.7	735	4	B1756228	B1756228 603024265
13	391.2	20.4	666	4	B1480798	B1480798 H2RPE-043
14	389.2	20.3	639	6	CA389968	CA389968 ca310403
15	381.2	19.9	491	7	CV024947	CV024947 2492 Full1
16	368.4	19.2	582	5	BP379588	BP379588 BP379588
17	367	19.2	508	2	BE216901	BE216901 145906 MA
18	342.2	17.9	526	4	BM685501	BM685501 UT-E-C10-
19	336	17.5	526	5	BM707649	BM707649 UT-E-C11-
20	334	17.4	522	5	BM707649	BM707649 UT-E-C11-
21	333.2	17.4	522	5	BM707649	BM707649 UT-E-C11-
22	327.6	17.1	819	7	CR408463	CR408463 CR408463
23	319.2	16.7	934	7	CF241451	CF241451 AGENCOURT
24	310.6	16.2	497	6	CA391418	CA391418 ca15a10.Y

25	309	16.1	863	5	BU205267	BU205267 603103847
26	306	16.0	793	4	BG961794	BG961794 602826439
27	295.6	15.4	1067	4	BM562042	BM562042 AGENCOURT
28	291	15.2	318	1	AV040568	AV040568 AV040568
29	288.8	15.1	566	4	BM718338	BM718338 UT-E-RO1-
30	281.8	14.7	310	1	AV204954	AV204954 AV204954
31	279.8	14.6	298	5	BX634624	BX634624 BX634624
32	277.2	14.5	721	7	CN427090	CN427090 170004177
33	272.6	14.2	455	1	AV043872	AV043872 AV043872
34	272	14.2	455	4	BG013943	BG013943 ILS-GN024
35	269.2	14.1	282	1	AV043522	AV043522 AV043522
36	261.6	13.7	784	7	CF783151	CF783151 AGENCOURT
37	261	13.6	732	4	BG965806	BG965806 602829577
38	259.8	13.6	603	1	AA205892	AA205892 zq47a05.x
39	259.2	13.5	654	2	BE679896	BE679896 d58a11.Y
40	256.4	13.4	780	5	BX920526	BX920526 BX920526
41	254	13.3	271	1	AA497726	AA497726 v172b03.x
42	252.8	13.2	883	5	BU523270	BU523270 AGENCOURT
43	241.6	12.6	941	5	BU514796	BU514796 AGENCOURT
44	240.6	12.6	514	4	B1343182	B1343182 371242 MA
45	239.8	12.5	569	5	BP220982	BP220982 BP220982

ALIGNMENTS

RESULT 1
AK006549
LOCUS
DEFINITION Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:1700030H21 product:vtc11form macular dystrophy 2 homolog (human), full insert sequence.

ACCESSION AK006549
VERSION AK006549.1 GI:12839710
KEYWORDS HTG; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 11042159
PUBMED 11042159

REFERENCE
AUTHORS Kono, H., Akiyama, J., Nishii, K., Kitanai, T., Tashtiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Watanabe, H., Sekiguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 11076861
PUBMED 11076861

REFERENCE
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research

Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 1699)
Adachi, Y., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H.,
AUTHORS

Group Phases I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563–573 (2002)
6 (bases 1 to 1699)
Aachari, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H.,
Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,
Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,
Hiraoaka, T., Hori, F., Imochi, K., Iishi, Y., Itoh, M., Izawa, M.,
Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M.,
Koyama, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishik, K.,
Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,
Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
Shibata, K., Shibata, Y., Shingawa, A., Shiraki, T., Sogabe, Y.,
Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T.,
Tejima, Y., Toyota, T., Yamamura, T., Yasunishi, A., Yoshida, K.,
Yoshino, M., Muramatsu, M., and Hayashizaki, Y.

TITLE Direct Submission
JOURNAL Submitted (10-JUL-2000) Yoshitake Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: <http://genome.gsc.riken.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT Please visit our web site (<http://genome.gsc.riken.jp/>) for further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Science Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GAGGAGAGAGAGATCTCAAGACCTCTTTTCTTTTCTTTTCTTTTCA 3'], cDNA was prepared by using triethylalco thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5'-GAGGAGAGAGAGGCGCCGCAATTATTCCTGAGTATTAATTAATACCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI, 3' end: SstI. Host: SOLR.

FEATURES	
source	Location/Qualifiers
	1. .1699
	/organism="Mus musculus"
	/mol_type="mRNA"
	/strain="C57BL/6J"
	/db_xref="PANTOM.DB:1700030H21"
	/db_xref="taxon:10090"
	/clone="1700030H21"
	/sex="male"
	/tissue_type="testis"
	/clone_lib="RIKEN full-length enriched mouse cDNA library"
	/dev_stage="adult"
	1. .1699
misc_feature	/note="Vitellogenin macular dystrophy 2 homolog (human) (MGD GI:1346332)"

Query Match	87.9%	Score 1684.8;	DB 3;	length 1699;
Best Local Similarity	99.8%;	Pred. No. 0;		
Matches 1697;	Conservative	0;	Mismatches 2;	Indels 1;
				Gaps 11

OY	179	GATCAGCAGCGTGTGTTGAGAGCGTGCTGTGACCTGACACGCTAACATTAAGCTCAGC	238
Db	1	GATCAGCAGCGTGTGTTGAGAGCGTGCTGTGACCTGACACGCTAACATTAAGCTCAGC	60
OY	239	CCTATATCCTTCGTTCTGGGTTTCTATGTTACATTTGGTGTGAGCCGCGTGAGCCAG	298
Db	61	CCTATATCCTTCGTTCTGGGTTTCTATGTTACATTTGGTGTGAGCCGCGTGAGCCAG	120
OY	299	TACGAGAACTTGCCTGGCCGACCGGCTCATGATCCAGGTGTCTAGCTTCGTGAGGGC	358
Db	121	TACGAGAACTTGCCTGGCCGACCGGCTCATGATCCAGGTGTCTAGCTTCGTGAGGGC	180
OY	359	AAGATGAGGAAGCGCTTTCCTCGCGGCGACGCTCATTCGTCACCCCATCTCGGGCCAA	418

Db	181	AAAGATGAAGAAAGCGCTTGTCTGGGCGCAGCGCTATCGCTACGCAATCTGGGCGAA	240
QY	419	GTGCTCATCTGCGAGCATCAAGCACTTGGTCTAACAGCGCTTTCACCTCTTCAACAC	478
Db	241	GTGCTCATCTGCGAGCATCAAGCACTTGGTCTAACAGCGCTTTCACCTCTTCAACAC	300
QY	479	CTGGTGCTAGCAGGTTTTATGACCCATGAGGAAACATTAAGCATGTGCAAAATTGGGCGCTA	538
Db	301	CTGGTGCTAGCAGGTTTTATGACCCATGAGGAAACATTAAGCATGTGCAAAATTGGGCGCTA	360
QY	539	CCACACAACACATTTGGGTGCGCTGGGTGTGGTTTGCCAACTTGCTCAATGAAGCGCTAT	598
Db	361	CCACA-AAACAATTGGGTGCGCTGGGTGTGGTTTGCCAACTTGCTCAATGAAGCGCTAT	419
QY	599	CTTGAAGGTCCGAATCCGGGACACCGGTCTGCTCCAGAGCTGATGAATGAGGTGTGTA	658
Db	420	CTTGAAGGTCCGAATCCGGGACACCGGTCTGCTCCAGAGCTGATGAATGAGGTGTGTA	479
QY	659	TTCGGTACTCAGTGTGGA CAGCTGTATCCCTACGATAGATTAAGTATCCATTTGGTGTGAC	718
Db	480	TTCGGTACTCAGTGTGGA CAGCTGTATCCCTACGATAGATTAAGTATCCATTTGGTGTGAC	539
QY	719	AACAAGTGTGTACAGTGGCAGGTATACAGCTTTTTCCTTGACGTGATCGGAGGCGAG	778
Db	540	AACAAGTGTGTGTACAGTGGCAGGTATACAGCTTTTTCCTTGACGTGATCGGAGGCGAG	599
QY	779	TTTCTGAAACCCAAACAAGGACTACCCAGGCGCATGATGATCTTGGTGTGGCTGTCTTC	838
Db	600	TTTCTGAAACCCAAACAAGGACTACCCAGGCGCATGATGATCTTGGTGTGGCTGTCTTC	659
QY	839	ACAATCTGTCAATTTCTTATTTCACTAGGCGTGGCTGAAGGTGGCAGAAACAGCTCATCAC	898
Db	660	ACAATCTGTCAATTTCTTATTTCACTAGGCGTGGCTGAAGGTGGCAGAAACAGCTCATCAC	719
QY	899	CCCTTCGGGGAGAGCATGATGATTTTGAAGCTAACATGATCATTTGA CAGAAACCTGCAG	958
Db	720	CCCTTCGGGGAGAGCATGATGATTTTGAAGCTAACATGATCATTTGA CAGAAACCTGCAG	779
QY	959	GTGTCCCTGTGTCCGTGGATGGGATGTGACAGAACTTGCTCCCATGGAAGTGA CATTG	1018
Db	780	GTGTCCCTGTGTCCGTGGATGGGATGTGACAGAACTTGCTCCCATGGAAGTGA CATTG	839
QY	1019	TACTGGAACGAGGCAAGCGCTCAGCCGCTTACACAGCTGTCTTTCGAGGTTTCGCGG	1078
Db	840	TACTGGAACGAGGCAAGCGCTCAGCCGCTTACACAGCTGTCTTTCGAGGTTTCGCGG	899
QY	1079	CATTCCTTCAGAGGCTCCACCTTCAACATGACCTTAAGAAAGAAAGAACTTGAAGCTTTGG	1138
Db	900	CATTCCTTCAGAGGCTCCACCTTCAACATGACCTTAAGAAAGAAAGAACTTGAAGCTTTGG	959
QY	1139	TCAAAAGAGAGGCGTGA CACGATTAAGAAAGAGTGCTTATAGCAGCAGCATAGGCTGC	1198
Db	960	TCAAAAGAGAGGCGTGA CACGATTAAGAAAGAGTGCTTATAGCAGCAGCATAGGCTGC	1019
QY	1199	TTCTTAGGACTGCAACCCAAACTTACATTTTCCTTGAAGACTTAAAGCCAAACTA	1258
Db	1020	TTCTTAGGACTGCAACCCAAACTTACATTTTCCTTGAAGACTTAAAGCCAAACTA	1079
QY	1259	TTGTGTTCTAAGAACCCCTCTCGAAGGCGCAGTGTAAAGATGCAAC CAGAAAC CAG	1318
Db	1080	TTGTGTTCTAAGAACCCCTCTCGAAGGCGCAGTGTAAAGATGCAAC CAGAAAC CAG	1139
QY	1319	AAAGATGTCTGGAATTTAAGGAGTGTGACTTCTTGAATGTGTTCAGAGTTTAAAGG	1378
Db	1140	AAAGATGTCTGGAATTTAAGGAGTGTGACTTCTTGAATGTGTTCAGAGTTTAAAGG	1199
QY	1379	AGAGGCTCCATTGTGGCCACAGGCA CCAAGCA CCACTTACTGAGCAAGTGA CAGCC	1438
Db	1200	AGAGGCTCCATTGTGGCCACAGGCA CCAAGCA CCACTTACTGAGCAAGTGA CAGCC	1259
QY	1439	TCCAGTTTCAGACAGGTGATGTGGCTTTCACAGATTTACAAAGAACTGTGCATGTAA	1498

Dh 1260 TCCAGTTGACACAGGTGATGGCGCTTCCACAGATTACCAAGAAATCTGTCACTAGAAA 1319
Qy 1499 AAGAAACTGTGAGATTAACTTGAACATTCACAGAGCCGCCACAGAACATCTTCAACAG 1558
Db 1320 AAGAAACTGTGAGATTAACTTGAACATTCACAGAGAGCCGCCACAGAACATCTTCAACAG 1379
Qy 1559 GCGCGTTTGAACAGATGTCAACCAATATACAGGCTTAATGAGAGACATGACAGAGTCC 1618
Db 1380 GCGCGTTTGAACAGATGTCAACCAATATACAGGCTTAATGAGAGACATGACAGAGTCC 1439
Qy 1619 TATCCCTACAGAGATGAGAGCTGGACCAACCTGTTCTATGAGTATGCTTCAACAGCC 1678
Db 1440 TATCCCTACAGAGATGAGAGCTGGACCAACCTGTTCTATGAGTATGCTTCAACAGCC 1499
Qy 1679 TGGCCCTGACTTGAACAGATGCCGACAGAGGCACTGACCCAGTCAAAAG3CACAAAGCAG 1728
Db 1500 TGGCCCTGACTTGAACAGATGCCGACAGAGGCACTGACCCAGTCAAAAG3CACAAAGCAG 1559
Qy 1739 CGACACCCAGAGAGTGTGTCCACAGACAGTCTAGCATGTAACTCAGAACCAAGAGTACTT 1798
Db 1560 CGACACCCAGAGAGTGTGTCCACAGACAGTCTAGCATGTAACTCAGAACCAAGAGTACTT 1619
Qy 1799 AATAGTCTGCTGAAAAACACTGTATTTTACGATCTTTCCCAAACCTAGAGATTAAATA 1858
Db 1620 AATAGTCTGCTGAAAAACACTGTATTTTACGATCTTTCCCAAACCTAGAGATTAAATA 1679
Qy 1859 AACGGAATATCTTTTAGG 1878
Db 1680 AACGGAATATCTTTTAGG 1699

RESULT 2
BO879880 963 bp mRNA linear EST 16-AUG-2002
LOCUS BO879880
DEFINITION AGENCOURT 8241331 Lupski_dorsal_root_ganglion Homo sapiens CDNA
clone IMAGE:6180559 5', mRNA sequence.
ACCESSION BO879880
VERSION BO879880.1 GI:22271888
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 963)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@dbp-remail.nih.gov
Tissue Procurement: Dr. James R. Lupski
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LAM1563 row: 1 column: 08
High quality sequence scop: 623.
Location/Qualifiers
1. .963
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6180559"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/clone_lib="Lupski_dorsal_root_ganglion"
/note="Vector: pCMV-Sport6 (Life Technologies); Site_1:
NotI; Site_2: SalI; CDNA made by oligo-CT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCACGCGTCCG-3' and

5'-GACTAGTTCTAGATCCGAGCGCCGCCCT(15)-3'. Size selected >
1 kb for average insert length 1.7 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies.

ORIGIN
Query Match 29.9%; Score 573.6; DB 5; Length 963;
Best Local Similarity 83.4%; Pred. No. 5,2e-139;
Matches 651; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

Qy 351 TGAAGGCGAAGATGAGAGAGCGGTTGCTGCGGCGCACGCTCATCGGATCC 410
Db 26 TGAAGGCGAAGATGAGAGAGCGGCTGCTGCGGCGCACGCTCATCGGATCC 85
Qy 411 TGGGCGAAGTCTATCTGCGGCGCATCAGACCTCGGTCTACAGAGGCTTTCCCATC 470
Db 86 TGGGCGAAGTCTATCTGCGGCGCATCAGACCTCGGTCTACAGAGGCTTTCCCATC 145
Qy 471 TTACACACCTGCTGAGAGGTTTATGACCCATGAGGAAATVAGAGTTCAGAGT 530
Db 146 CCGACACCTGCTGAGAGGTTTATGACCTCGGCGAAACAGAGAGTTGAGAAAC 205
Qy 531 TGGGCTTACACACACATTCCTGCGGCTCGGCTGCTTCCCACTGTCAATGA 590
Db 206 TGAGCTTACACACACATTCCTGCGGCTCGGCTGCTTCCCACTGTCAATGA 265
Qy 591 AGGCTTATCTTGGAGGTCGAATCCGGGACACCTGCTCTCAAGACCTGATGAATGAG 650
Db 266 AGGCTTATCTTGGAGGTCGAATCCGGGACACCTGCTCTCAAGACCTGCTGAACGAGA 325
Qy 651 TGTGTACTTTCGACTCAGTGTGAGACAGCTGTATGCTAGCATGATTAATCCCAT 710
Db 326 TGAACACCTTTCGACTCAGTGTGAGACAGCTGTATGCTAGCATGATTAATCCCAT 385
Qy 711 TGTGTACACACAGAGTGTGAGACAGTGTATGAGTTCCTTCTGATGCTTGAATCG 770
Db 386 TGTGTATACACAGAGTGTGAGACAGTGTATGAGTTCCTTCTGATGCTTGAATCG 445
Qy 771 GGAGGAGTTTCTGAACCCAAACAGAGCTTACCCAGGCTGATGATGATCGTGTGTC 830
Db 446 GGCGGAGTTTCTGAACCCAAACAGAGCTTACCCAGGCTGATGATGATGATGATGTC 505
Qy 831 CTGCTTCAACAATCTCTGAATCTTATCTTATGATGAGGCTGAGTGAAGTGCAGAACGC 890
Db 506 CCGTCTTCAACAATCTCTGAATCTTATCTTATGATGAGGCTGAGTGAAGTGCAGAACGC 565
Qy 891 TCATCAACCCCTTGGGAGAGAGCATGATGATTTTGAAGTATGATGATGATGATGATGATG 950
Db 566 TCATCAACCCCTTGGGAGAGAGCATGATGATTTTGAAGTATGATGATGATGATGATGATG 625
Qy 951 ACTGCAAGTGTCTCTGTTGCTGAGTGAATGATGACACAGAACTTGTCTCCATGAGAC 1010
Db 626 ATTGCAAGTGTCTCTGTTGCTGAGTGAATGATGACACAGAACTTGTCTCCATGAGAC 685
Qy 1011 GTGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1070
Db 686 CGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 745
Qy 1071 CTCGCGGAGATTCCTTCAAGAGGCTTCAACCTTCAACATCAGGCTTAAGAAAGAGACTTAG 1130
Db 746 TCCGCTGAGGCTCTTATGAGGCTTCAACCTTCAACATCAGGCTTAAGAAAGAGAGTGG 805
Qy 1131 A 1131
Db 806 A 806

RESULT 3
BY714331 713 bp mRNA linear EST 17-DEC-2002
LOCUS BY714331
DEFINITION BY714331 RIKEN full-length enriched, adult male testis Mus musculus

ACCESSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS
CU9A.G106	492150418	5'	mRNA	sequence.	
BY714331	1	GI:27127441			
EST.					
Mus musculus					
Mus musculus					
Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi, Mammalia, Eutheria, Rodentia, Sciurognathi, Muridae, Murinae, Mus					
1 (bases 1 to 713)					
Okazaki, Y., Furuno, M., Kasukawa, T., Aachi, J., Bono, H., Kondo, S., Nishida, I., Osato, N., Saito, R., Suzuki, H., Yamana, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schinohara, C., Gojobori, T., Baldorelli, R., Hill, D. P., Bule, C., Hune, D. A., Quackenbush, J., Schriml, L. M., Kapaplan, A., Matsuda, H., Batalov, S., Betsel, K. W., Blake, J. A., Brad, D., Brusic, V. A., Chochoa, C., Corbani, L. E., Cousins, S., Dalla, E., Dragani, T. A., Fletcher, C. F., Forrest, A., Fraser, K. S., Gaasterland, T., Gariboldi, M., Gissi, C., Godkij, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kani, A., Kawai, H., Kawasawa, Y., Kedzierski, R. M., King, B. L., Kondagala, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, M. J., Petrea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring, B. Z., Ringwald, M., Sandelin, A., Schneider, C., Semp, M. C., Secot, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tonita, M., Verdado, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L. G., Wynshaw-Boris, A., Yangisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kitahara, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shitaki, T., Waki, K., Kawai, J., Aizawa, K., Akakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kageawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shingawa, J., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S., Rogers, J., Birney, E. and Hayashizaki, Y.					
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs					
Nature 420, 563-573 (2002)					
22354683					
12468851					
COMMENT					
JOURNAL					
MECHANISM					
PUBMED					
CONTACT: Yoshihide Hayashizaki					
Laboratory for Genome Exploration Research Group, RIKEN Genomic					
Sciences Center (GSC), Yokohama Institute					
The Institute of Physical and Chemical Research (RIKEN)					
1-7-22 Suehiro-cho					
Tel.: 81-45-501-9222					
Fax: 81-45-501-9216					
Email: genome-res@gsc.riken.jp, URL: http://genome-gsc.riken.jp/					
Adachi, J., Aizawa, K., Akimura, T., Akakawa, T., Carninci, P., Fukuda, S., Hashizume, W., Hayashida, K., Hirozane, T., Horii, P., Imotani, K., Ishii, Y., Itoh, M., Kageawa, I., Kawai, J., Kojima, Y., Komura, S., Konno, H., Koyama, S., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Osato, N., Saito, R., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shitaki, T., Tagami, M., Takeda, Y., Waki, K., Watabiki, A., Muramatsu, M. and Hayashizaki, Y.					
Direct Submission					
Computational Analysis of Full-Length Mouse cDNAs Compared with					
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)					
Normalization and subtraction of cap-trapper-selected cDNAs to					
prepare full-length cDNA libraries for rapid discovery of new					
genes. Genome Res. 10 (10), 1617-1630 (2000)					
RIKEN integrated sequence analysis (RISA) system-384-format					
sequencing pipeline with 384 multicapillary sequencer. Genome Res.					
10 (11), 1757-1771 (2000)					
Computer-based methods for the mouse full-length cDNA					
encyclopedia: real-time sequence clustering for construction of a					
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)					
cDNA library was prepared and sequenced in mouse Genome					
Encyclopedia Project of Genome Exploration Research Group in Riken					
Genomic Sciences Center and Genome Science Laboratory in RIKEN.					
Division of Experimental Animal Research in Riken contributed to					
prepare mouse tissues.					

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES	Location/Qualifiers
source	1..713

ORIGIN

/clone_id=Riken full length enriched, adult male testis /site=1: XhoI, site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGGAGAGAGAGATCCAGAGCTCTTTTCTTTTTTTTTTNN 3']. cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adaptor of sequence [5', GAGGAGAGAGAGCGCGCCATTAATTCCTGAGTTAATTAATATATATCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. "

Query Match	27.1%	Score 519;	DB 6;	Length 713;
Best Local Similarity	97.8%;	Pred. No. 1e-124;		
Matches 525; Conservative	0;	Mismatches 12;	Indels 0;	Gaps 0;

QY	GTGGCAAGCCATATACATATACCTACACAAACAAATGACCATGCGGCTCGGTTTCGTT	60
Db	GTGGCAAGCCATATACATATACCTACACAAACAAATGACCATGCGGCTCGGTTTCGTT	120
QY	CTGTGCTCTCTCTCTGTGTGGCGGAGGAGCATCTACAAAGCTGTGTATGGAGAAATTCCT	120
Db	CTGTGCTCTCTCTCTGTGTGGCGGAGGAGCATCTACAAAGCTGTGTATGGAGAAATTCCT	236
QY	TGTCTTCATATTCCTCTACTATTTCCATCCGTGGACTCTACAGAAATGGTTCTCTCGAGTGA	180
Db	TGTCTTCATATTCCTCTACTATTTCCATCCGTGGACTCTACAGAAATGGTTCTCTCGAGTGA	356
QY	TCAGCAGCTGTTGTTTGAGAAAGCTGGCTCTGTACTGTGGACAGCTACATTCAGCTCATCCC	240
Db	TCAGCAGCTGTTGTTTGAGAAAGCTGGCTCTGTACTGTGGACAGCTACATTCAGCTCATCCC	416
QY	TATATCTCTCGTTCTGGGGTTTCTAATGTTAACTTGGTGTGTAGCCGCTGTGTGAGCCAGTA	300
Db	TATATCTCTCGTTCTGGGGTTTCTAATGTTAACTTGGTGTGTAGCCGCTGTGTGAGCCAGTA	476
QY	CGAAGACTTGGCGGTGAGCCGACCGGCTCATGATCAGAGTGTCTTAGCTCTGTGGAGGGCAA	360
Db	CGAAGACTTGGCGGTGAGCCGACCGGCTCATGATCAGAGTGTCTTAGCTCTGTGGAGGGCAA	536
QY	GGATGAGGAAGGCGGTTTGTGTGGGGGCGCAAGCTCATCCGTTACGCCATCTCTGGGCCAAGT	420
Db	GGATGAGGAAGGCGGTTTGTGTGGGGGCGCAAGCTCATCCGTTACGCCATCTCTGGGCCAAGT	596
QY	GCTCATCTGTGGCAGCATCAGCACTTGAGTCTCAAGAGCGTTTCCCACTTTCACCAACT	480
Db	GCTCATCTGTGGCAGCATCAGCACTTGAGTCTCAAGAGCGTTTCCCACTTTCACCAACT	656
QY	GGTGTACAGAGTTTATAGCCATGGGGAACTAAGCAAGTTGACAAATTTGGAGCT	537
Db	GGTGTACAGAGTTTATAGCCATGGGGAACTAAGCAAGTTGACAAAGTTTGGAGCT	713

RESULT 4
CD518675

LOCUS CD518675 881 bp mRNA linear EST 06-JUN-2003
 DEFINITION AGNCOURT 14375996 NIH_MGC_181 Homo sapiens cDNA clone
 IMAGE:3037964 5', mRNA sequence.
 ACCESSION CD518675
 VERSION CD518675.1 GI:31450393
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 881)
 NIH-MGC http://mgi.nci.nih.gov/
 ADHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL Contact: Daniela S. Gerhard, Ph.D.
 COMMENT Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: Dr. Michael Brownstein
 cDNA Library Preparation: Invitrogen Corp
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: NDAM482 row: b column: 21
 High quality sequence stop: 652.
 Location/Qualifiers
 1..881
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3037964"
 /tissue_type="White Matter"
 /dev_stage="Unknown"
 /lab_host="DH10B-Ton A (T1 and T5 phage resistances)"
 /clone_1lb="NIH_MGC_181"
 /note="Vector: pCMV-SPORT6.1; Site_1: NotI; Site_2: EcoRV
 (destroyed); Library is oligo-dT primed and directionally
 cloned (EcoRV site is destroyed upon cloning). Average
 insert size 1.42 kb. Library was constructed by
 (Invitrogen). Note: this is a NIH_MGC Library."

ORIGIN
 Query Match 26.2%; Score 502.6; DB 6; Length 881;
 Best Local Similarity 80.1%; Pred. No. 2.2e-120;
 Matches 608; Conservative 0; Mismatches 139; Indels 12; Gaps 1;

QY 476 CACCTGGTGTGACGAGTTTATGACCGATGCGGAGCAATAGACAGTTCCAGAGTTGGGC 535
 Db 1 CACCTGGTGTGACGAGTTTATGACCGATGCGGAGCAATAGACAGTTCCAGAGTTGGGC 60

QY 536 CTACACACACACATTTCTGGGTGCTGCTGGGTGCTGGTGCACATTTCTCATAGAGGCC 595
 Db 61 CTACACACACACATTTCTGGGTGCTGCTGGGTGCTGGTGCACATTTCTCATAGAGGCC 120

QY 596 TATCTGGAGGTGCAATCCGGGACACCGTCTGCTCCAGAGCTGATTAATAGAGTGTGT 655
 Db 121 TGGCTTGGAGGTGCAATCCGGGACACCGTCTGCTCCAGAGCTGATTAATAGAGTGTGT 180

QY 656 ACTTTGGCTACTGACGTGTGACAGCTGTATGCTGATGCTGATTAATAGATCCATTTGGTG 715
 Db 181 ACTTTGGCTACTGACGTGTGACAGCTGTATGCTGATGCTGATTAATAGATCCATTTGGTG 240

QY 716 TACACACAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGT 775
 Db 241 TATACACAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGT 300

QY 776 CAGTTTCTGAACCCCAAGAGCTAGCCAGGCCATGAGATGATCTGCTGTGCTGCTGTC 835
 Db 301 CAGTTTCTGAACCCCAAGAGGCCCTAGCCCTGAGCTGAGTGAAGTCTGCTGTGCTGCTGTC 360

QY 836 TTCAACATCTCTGCAATTTCTTATTTACATAGGCTGCTGAGAGTGGCAGAAACAGTCTATC 895
 Db 361 TTCAAGTTCCTGCAAGTTCTTTCTATATGTGGCTGGCTGAGAGTGGCAGAGCTCTATC 420

QY 896 AACCCCTTGGGAGAGACGATGATGATTTTGAAGTAAGTGAATATTGACAGAAACCTG 955
 Db 421 AACCCCTTGGGAGAGATGATGATGATTTTGAAGTAAGTGAATATTGACAGAAATTTG 480

QY 956 CAGGTGCTGCTGTTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1015
 Db 481 CAGGTGCTGCTGTTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 540

QY 1016 ATGTACTGGAACGAGGACGCTGACGCTGACGCTGACGCTGCTGCTGCTGCTGCTGCTG 1075
 Db 541 ATGTACTGGAATTAAGCCGACGACGACGACGACGACGACGACGACGACGACGACGACG 600

QY 1076 CGGACTTCTTCAATGAGCTCCACCTTCAACATCAGCTTAAAGAAAGAAAGACTTAGACTT 1135
 Db 601 CGAGCTTCTTCAATGAGCTCCACCTTCAACATCAGCTTAAAGAAAGAAAGACTTAGACTT 660

QY 1136 TGTCTAAGAGAGAGGCTGACACGATTAAGAAAGAGTGTGCTATAGACAGCACAAGAG 1195
 Db 661 CAGCCCAATCAGAGAG-----ACAGAGAGATGCTCAGCTGGCATCTTTGAC 708

QY 1196 TGTCTTGTAGAGCTGCAACCCAAACCTACCTATCTTCCC 1234
 Db 709 CGTTTCTTGTAGAGCTGCAACCTACCTATCTTCCC 747

RESULT 5
 CK470840 792 bp mRNA linear EST 13-JAN-2004
 LOCUS AGNCOURT 17639671 NIH_MGC_237 Rattus norvegicus cDNA clone
 DEFINITION IMAGE:7115295 5', mRNA sequence.
 ACCESSION CK470840
 VERSION CK470840.1 GI:40809322
 KEYWORDS EST.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 792)
 NIH-MGC http://mgi.nci.nih.gov/
 ADHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL Contact: Daniela S. Gerhard, Ph.D.
 COMMENT Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: Howard Jacobs
 cDNA Library Preparation: Express Genomics
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LRAM1490 row: d column: 13
 High quality sequence stop: 648.
 Location/Qualifiers
 1..792
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 /clone="IMAGE:7115295"
 /tissue_type="testis, pooled"
 /lab_host="DH10B-TonA"
 /clone_1lb="NIH_MGC_237"
 /note="Organ: testis; Vector: pExpress-1; Site_1: EcoRV;
 Site_2: NotI; RNA obtained from testis tissue of 8 wk old
 animal. Tissues were snap-frozen and kept at -80C before
 RNA extraction and purification (TRI-reagent method). cDNA

was primed using oligo-dt primer: 5'-pGACTATCTTTAATGCGGAGCGCGCC (75-3', and cloned into the EcoRV/NotI sites of pExpress-1. Size-selection >1.4Kb resulted in an average insert size of 2.4 Kb. This primary library is not normalized (normalized primary library is NIH-McG_28) and was constructed by Express Genomics (Frederick, MD)"

QY	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS
963	BT14344	664 bp mRNA linear EST 17-DEC-2002	BT14344						
548	BT14344	RIKEN full-length enriched, adult male testis Mus musculus	BT14344						
923	TTTGAGACTAATCGATCATTTGACAGAAACCTGACGATGCTCCCTGTTCCGCTG	976							
608	TTTGAGACCAATTGGATTGTGCAAGGAATTGGCAAGTGCTCCCTGTTGCTGTG	661							
963	BT14344	664 bp mRNA linear EST 17-DEC-2002	BT14344						
548	BT14344	RIKEN full-length enriched, adult male testis Mus musculus	BT14344						
923	TTTGAGACTAATCGATCATTTGACAGAAACCTGACGATGCTCCCTGTTCCGCTG	976							
608	TTTGAGACCAATTGGATTGTGCAAGGAATTGGCAAGTGCTCCCTGTTGCTGTG	661							

Query Match	Best Local Similarity	25.4%; Score 487; DB 6; Length 664;
Matches	487; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
QY	1	GTGCCAAGCCATGACATACCTACCTACACAAACAAAGTAGCCATGCCCGCTCGGTTGTT 60
DB	177	GTGCCAAGCCATGACATACCTACCTACACAAACAAAGTAGCCATGCCCGCTCGGTTGTT 236
QY	61	CTGCTCCCTCCTCGTGCTGGGAGGAGAGATACACAGCTGCTGTATGAGAAATTCCT 120
DB	237	CTGCTCCCTCCTCGTGCTGGGAGGAGAGATACACAGCTGCTGTATGAGAAATTCCT 296
QY	121	TGCTTCATATTCCTCTACTATTTCATCCATCCGTGACCTTACAGATGGTTCTTCGATGA 180
DB	297	TGCTTCATATTCCTCTACTATTTCATCCATCCGTGACCTTACAGATGGTTCTTCGATGA 356
QY	181	TCAGACAGCTGTTGTTGAGAGAGCTGCTCTGTACTGCGACAGACTACATTCACTCATCC 240
DB	357	TCAGACAGCTGTTGTTGAGAGAGCTGCTCTGTACTGCGACAGACTACATTCACTCATCC 416
QY	241	TATATCCCTGCTTGGGGTTTCTATGTTAATGATGGTGTAGACCGCGGTGGAGCCATGA 300
DB	417	TATATCCCTGCTTGGGGTTTCTATGTTAATGATGGTGTAGACCGCGGTGGAGCCATGA 476
QY	301	CGAAGACTTGGCCGTGGCCGACCGCTCATGATCCAGGTCTTAGCTTCGTGAGAGGCAA 360
DB	477	CGAAGACTTGGCCGTGGCCGACCGCTCATGATCCAGGTCTTAGCTTCGTGAGAGGCAA 536
QY	361	GGATGAGGAAGGCGTTGCTGGGGCCGACGCTCATCCGCTAGGCCATCTCGGCGCAAGT 420

Db 537 GGATGAGAGAGCGCTTTGCTGCGGCGACGCTATCCGCTACGCCATCTCGGCCCACT 596
 Qy 421 GGTATCTCGCGGACATCAGCACTCGGTCTACAGAGCGCTTTCCACTTTCCACCACT 480
 Db 557 GGTATCTCGCGGACATCAGCACTCGGTCTACAGAGCGCTTTCCACTTTCCACCACT 656
 Qy 481 GGTCCTAG 488
 Db 657 GGTGCTAG 664
 RESULT 8
 LOCUS CK469826 862 bp mRNA linear EST 13-JAN-2004
 DEFINITION AGNCOURT 17644559 NIH MGC 237 Rattus norvegicus cDNA clone
 IMAGE:7113326 5', mRNA sequence.
 ACCESSION CK469826
 VERSION CK469826.1 GI:40808308
 KEYWORDS EST.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 862)
 AUTHORS NIH-MGC http://mgs.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Daniela S. Gerhard, Ph.D.
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cgadbs-remail.nih.gov
 Tissue Procurement: Howard Jacobs
 CDNA Library Preparation: Express Genomics
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILMIN)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/ILMIN at:
 http://image.llnl.gov
 Place: L1M41985 row: b column: 12
 High quality sequence stop: 674.
 Location/Qualifiers
 1..862
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 /clone="IMAGE:7113326"
 /tissue_type="testis, pooled"
 /lab_host="DH10B Tona"
 /clone_lib="NIH MGC 237"
 /note="Organ: testis; Vector: pExpress-1; Site_1: EcoRV;
 Site_2: NotI; RNA obtained from testis tissue of 8 wk old
 animal. Tissues were snap-frozen and kept at -80C before
 RNA extraction and purification (TRI-reagent method). cDNA
 was primed using oligo-dT primer:
 5'-GGACTAGTTCTAGATCGGAGCGCGCCGCTT(25-3' and cloned into
 the EcoRV/NotI sites of pExpress-1. Size-selection >1.4kb
 resulted in an average insert size of 2.4 kb. This primary
 library is not normalized (normalized primary library is
 NIH-MGC 238) and was constructed by Express Genomics
 (Frederick, MD)"

ORIGIN
 Query Match 24.7%; Score 474.2; DB 7; Length 862;
 Best Local Similarity 92.0%; Pred. No. 6.3e-113;
 Matches 532; Conservative 0; Mismatches 41; Indels 5; Gaps 3;

Qy 1 GTGCAAGCCATGATCATCACTACCAACAAAGTAGCAATGCCGCTGGTTCGT 60
 Db 158 GTGCAAGCCATGATCATCACTACCAACAAAGTAGCAATGCCGCTGGTTCGT 217
 Qy 61 CTGCTCCTCTCTCTGCTGCTGCGAGGAGCATCTCAAGCTGCTGTATGAGAAATTCCT 120

Db 218 CTGAGGCTCTCTCGGCGCTGGCGAGGACATCTCAAGCTGTGACGAGAAATTCCT 277
 Qy 121 TGTCTTCATATTTCTCTACATATTCATCCGTGACCTCTACAGATGTTCTCTGAGTGA 180
 Db 278 TGTCTTCATATTTCTCTCTACTATTCATCCGTGACCTCTACAGATGTTCTCTGAGTGA 337
 Qy 181 TCAGCAGCTGTGTTGTTGAGAGCTGCGTCTGACTGTGAGACAGTACTTACGCTCATCCC 240
 Db 338 TCAGCAGCAGCAGCTTTGAGAGATTGGCTCTGACTGTGAGACAGTACTTACGCTCATCCC 397
 Qy 241 TATATCTCTGTTCTGCGTTTCTATATTTACATTTGTTGAGAGCGCTGTGAGCAGTA 300
 Db 398 TATATCTCTGTTCTGCGTTTCTATATTTACATTTGTTGAGAGCGCTGTGAGCAGTA 457
 Qy 301 CGAAGACTTGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
 Db 458 CGAAGACTTGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 517
 Qy 361 GGATGAGAGAGCGGCTTTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420
 Db 518 GGAAGAGAGAGCGGCTTTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 577
 Qy 421 GGTATCTCTGCGGAGATACGACCTTCGTTCTACAGAGCGCTTTCCACTTTCCACCACT 480
 Db 578 GGTATCTCTGCGGAGATACGACCTTCGTTCTACAGAGCGCTTTCCACTTTCCACCACT 637
 Qy 481 GGTCCTAGCAGTTTATGATACCACTGGGAGACATAGCAGTGGAGAGTTGGGCTTACC 540
 Db 638 GATGCAAGCAGGCTTTATGATCTCTGNGAATTA-AGCAGTTGCAAGAGTGGGCTTACC 696
 Qy 541 ACACACACATTTTGGGTGCGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 578
 Db 697 AC--ACACATCTGGGT-ACCTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 730

RESULT 9
 LOCUS AU040308/c 578 bp mRNA linear EST 04-DEC-1998
 DEFINITION AU040308 Mouse four-cell-embryo cDNA Mus musculus cDNA clone
 J0803G03 3', mRNA sequence.
 ACCESSION AU040308
 VERSION AU040308.1 GI:3954128
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 578)
 AUTHORS Ko, M.S.H., Kitchen, J.R., Wang, X., Threat, T.A., Sun, T.,
 DePalma, G.E., Liang, Y., Kargul, G.J., Shatara, R., Lim, M.K. and
 Doi, H.
 TITLE Systematic analyses of genes expressed in 4-cell mouse embryo (The
 ERAO/Doi Project at Wayne State University)
 JOURNAL Unpublished (1998)
 COMMENT Contact: Hirofumi Doi
 Doi Biosymmetry Project, ERAO
 Japan Science and Technology Corporation (JST)
 WBG Marine East 12F, 2-6 Nakase, Mihama-Ku, Chiba 261-71, Japan
 Email: hdoibioa.jst.go.jp.
 Location/Qualifiers
 1..578
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="J0803G03"
 /dev_stage="four-cell-embryo"
 /clone_lib="Mouse four-cell-embryo cDNA"

ORIGIN

Query Match 24.3%; Score 466; DB 1; Length 578;
 Best Local Similarity 94.6%; Pred. No. 8.2e-111;

CDNA Library Preparation: Express Genomics
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: L1AM14988 row: 9 column: 05
High quality sequence stop: 669.
Location/Qualifiers

FEATURES

source

```
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="IMAGE:7114591"
/cisue_type="testis, pooled"
/lab_host="DH10B Tona"
/clone_lib="NIH_MGC_237"
/notes="Organ: testis; Vector: pExpress-1; Site_1: EcoRV;
Site_2: NotI; RNA obtained from testis tissue of 8 wk old
animal. Tissues were snap-frozen and kept at -80C before
RNA extraction and purification (Tri-reagent method). CDNA
was primed using oligo-dT primer:
5'-pGACTGATCTAGATGCGAGCGGCGCCCT(7)25-3' and cloned into
the EcoRV/NotI sites of pExpress-1. Size-selection >1.4kb
resulted in an average insert size of 2.4 kb. This primary
library is not normalized (normalized primary library is
NIH_MGC_238) and was constructed by Express Genomics
(Frederick, MD)"
```

ORIGIN

Query Match

Best Local Similarity 22.3%; Score 428.2; DB 7; Length 844;

Matches 442; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

```

1 GTGCCAAGCATGATGATACCTTACACAAACAAAGTAGCAATGCCGCTGGTGGTT
182 GTGCCAAGCATGATGATACCTTACACAAACAAAGTAGCAATGCCGCTGGTGGTT
61 CTCGTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
242 CTCGTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
121 TGTCTTCATATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
302 TGTCTTCATATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
181 TCAGCAGCTGTTGTTGAGAGCTGCTCTGACTGCGACAGCTACATTCAGTCATCCC
362 TCAGCAGCAGCAGTTGAGAGTTGGCTCTGACTGCGACAGCTACATTCAGTCATCCC
241 TATATCCTTCGTTCTGGGTTCTATGTTACATTTGGTGTGAGCCGCTGGTGGAGCA
422 TATATCCTTCGTTCTGGGTTCTATGTTACATTTGGTGTGAGCCGATGGAGCAAGTA
301 CGAGAACCTTGGCGTGGCCGACCGCTCATGATCAGAGTGTCTGAGTTCGAGAGGCA
482 CGAGAACCTTGGCGTGGCCGACCGCTCATGATCAGAGTGTCTGAGTTCGAGAGGCA
361 GGAGTAGAGAGAGCGGTTGCTGCGGCGACCGCTCATCAGCTACCGCATTCCTGGGCA
542 GGAAGAGAGAGAGCGGTTGCTGCGGCGACCGCTCATCAGCTACCGCATTCCTGGGCA
421 GCTCATCTGGGCGAGCATCAGCAGCTGGTGTATTAAGCGCTTTTC 465
602 GCTCATCTGGGCGAGCATCAGCAGCTGGTGTATTAAGCGCTTTTC 646
```

RESULT 12
BI756228 735 bp mRNA linear EST 25-SEP-2001
LOCUS 603024265F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5194649 5',
DEFINITION mRNA sequence.
ACCESSION BI756228

VERSION

BI756228.1 GI:15747806

EST.

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 735)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-rc@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:

<http://image.llnl.gov>

Plate: L1AM1487 row: a column: 18

High quality sequence stop: 577.

Location/Qualifiers

1..735

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5194649"

/lab_host="DH10B"

/clone_lib="NIH_MGC_114"

/notes="Organ: Brain; Vector: PCMV-SPORT6; Site_1: NotI;

Site_2: EcoRV (destroyed); RNA source anonymous pool of 6

male brains; age range 23-27 yo. Library is oligo-dT

primed and directionally cloned (EcoRV site is destroyed

upon cloning). Average insert size 1.5 kb, insert size

range 1-3 kb. Library is normalized and enriched for

full-length clones and was constructed by C. Gruber

(Invitrogen). Research Genetics tracking code 019. Note:

this is a NIH_MGC Library."

ORIGIN

Query Match

Best Local Similarity 20.7%; Score 397.4; DB 4; Length 735;

Matches 474; Conservative 0; Mismatches 91; Indels 2; Gaps 2;

```

144 CCATCCGTAAGCTCTGACAGAAAGTGTCTCTGACAGTATGAGCAGGTTGTTGAGAGC
40 CCACACTGGAACCCACCTGCTGGCCCTTACGAGAAACAGCTGATGTTGAGAAC
99
204 TGGCTCTGTAATGCGACAGCTACATTCAGCTCATCCATATCTTCTGTTGGGTTCT
100 TGACTCTGTAATGCGACAGCTACATTCAGCTCATCCCATTTCTTCTGTTGGGTTCT
264 ATGTTACATTTGATGAGAGCCGCTGTGAGAGCCATTAAGAACTTGGCCGCGAC
160 AGCTAGAGCTGATGCTGAGACCGCTGTGGAACAGTATGAGAACTTGGCCGCGAC
219
324 GCTCATGATGATGAGTGTCTGATGCTGAGAGGCAAGATGAGAAAGCCGTTTGGC
383
220 GCTCATGAGCCTGATGTTGAGGCTTCTGTCGAGAGCAAGAGAGAGAGCCGCTGCTC
279
384 GGGCAGCGCTCATCCGCTACCGCATCTGGGCGCAAGTCTCATCTCTGGCAGATCACA
443
280 GGGCAGCGCTCATCCGCTACCGCATCTGGGCGCAAGTCTCATCTCTGGCAGATCACA
339
444 CCTCGATGATGAGAGGCTTCCCACTCTTCCACACCTGTGAGTGTGAGAGGTTTATGACC
503
340 CCGCAGCTTACAGAGGCTTCCCACTGTGAGAGGCTTGTGAGAGGCTTGTGAGTCTC
399
504 ATGGGAAACATTAAGCAGTTTGAAGAGTTGGCTTACACACATTCCTGGTGGCTCT
563
400 CGGACAGAACACAGAGAGTTGAGAACTGAGACCTTACACACATTCCTGGTGGCTCT
459
564 GGGTGTGTTGGCCAACTTGTCAATGAAGGCTTATCTTGGAGGTGCAATCCGGGACACCG
623
```

Db 460 GGGTGTGGTTGGCAACCTGTCATGAAGGCGTGGCTTGGAGTGAGATCCGGGACCCCTA 519
Qy 624 TCCCTGCTCCAAAGCCTGATGAATGAGTGTGTA-CTTGGCTACTCAATGTGGACAGCTG 682
Db 520 TCCGCTCCAAAGCCTGCTGAACGAGATGAACACCTTGGCGTTCGGTGTGGACACCTG 579
Qy 683 TATGCGCT-ACGACTGATAGTATGCC 708
Db 580 TATGCTTAACGACTGATAGTATGCC 606

RESULT 13
BI480798 666 bp mRNA linear EST 28-FEB-2002
LOCUS H2PPE-0430 Human Retinal Pigment Epithelium (2) Homo sapiens cDNA
DEFINITION 5' similar to vitellogenin macular dystrophy (best disease, mRNA
sequence.
VERSION BI480798
KEYWORDS EST. BI480798.1 GI:18998607
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Buraczynska,M., Mears,A.J., Zarepari,S., Farjo,R., Filipova,E.,
AUTHORS Yuan,Y., Macnee,S.P., Hughes,B. and Swaroop,A.
TITLE Towards an expression profile of native human retinal pigment
epithelium: Identification of a non-redundant set of more than 1100
genes
JOURNAL Unpublished (2001)
COMMENT Contact: Swaroop, A.
Department of Ophthalmology and Visual Sciences
Kelllogg Eye Center, University of Michigan
540 KEC, 1000 Mail St., Ann Arbor, MI 48105, USA
Tel: 734 615 2246
Fax: 734 647 0228
Email: swaroop@umich.edu
PCR Primers
FORWARD: M13/PUC-Reverse - cccagtcacgacgtcgttaaacg
BACKWARD: M13/PUC-Forward - agcgagatacaatcctcacacg
Seq primer: M13/PUC-Reverse.
FEATURES
source
1..666
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="Native Retinal Pigment Epithelium sheets"
/dev_stage="juvenile"
/clone_lib="Human Retinal Pigment Epithelium (2)"
/note="Organ: Retina; Vector: pSport1"

ORIGIN
Query Match 20.4%; Score 391.2; DB 4; Length 666;
Best Local Similarity 81.4%; Pred. No. 3.7e-91;
Matches 464; Conservative 0; Mismatches 105; Indels 1; Gaps 1;

Qy 2 TGGCAAGCATGACTATACCTACCAACAAAGTAGCAATGCCGCTCGGTTCGTTTC 61
Db 98 TGGCGGCGCATGACCATCACTTACCAAGCCAGTGAATGCCGCTTAGGCTCCCTTC 157
Qy 62 TCGTCCCTCCCTGCTGCTGGCGAGGAGCATCTTCAAGCTGCTGTATGAGAAATTCCTT 121
Db 158 TCCCGCCCTGCTGCTGCTGGCGGAGCATCTTCAAGCTGCTGTATGAGAGTTCTTA 217
Qy 122 GTCTTCATATTCCTTACTATTTCCATCCGTGAGACTCTCAGAAATGTTCTTCGAGAT 181
Db 218 ATCTTCTGCTGCTGCTACTACATCATCGGCTTTATTTATAGGCTGAGCCCTTACGAGAA 277
Qy 182 CAGCAGCTGTGTTTGAAGAGCTGGCTGTACTGCGACAGCTACATTCACTCATTCCT 241
Db 278 CACAGCTGATGTTTGAAGAACTGACTCTGTATTCGACAGCTACATCCAGCTCATCC 337

Qy 242 ATATCTTCGTTCTGGGTTTCTATATGTTACATTTGGTGGAGACCCGCTGGTGGAGCCAGTAC 301
Db 338 ATTTCTTCCTGCTGCTGGGCTTTACGTAGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 397
Qy 302 GAGAACTTGGCCGCGCGGACCGGCTCATGATGATGATGATGATGATGATGATGATGATGATG 361
Db 398 GAGAACTGCGCGGCGCGGACCGGCTCATGATGATGATGATGATGATGATGATGATGATGATG 457
Qy 362 GATAGAGAGGCGCTTTGCTGCGGCGGACCGCTCATGATGATGATGATGATGATGATGATGATG 421
Db 458 GACAGAGAGGCGCGCTGCGGCGGACCGCTCATGATGATGATGATGATGATGATGATGATGATG 517
Qy 422 CTCATCTGCGGACGATCAGACCTCGGCTTACAAAGGCTTCCACTTCCACCACTG 481
Db 518 CTCATCTGCGGACGATCAGACCTCGGCTTACAAAGGCTTCCACTTCCACCACTG 576
Qy 482 GTGCTAGAGGTTTATGACCCATGGGAGAACATGAGTTCAGAAAGTTGGAGGCTTACCA 541
Db 577 GTGCAAGAGGCTTTATGACTTACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 636
Qy 542 CACAACATTTCTGGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCG 571
Db 637 CACAACATTTCTGGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCG 666

RESULT 14
CA389968 639 bp mRNA linear EST 06-NOV-2002
LOCUS ce104h03.y1 Human Retinal Pigment Epithelium/choroid cDNA
DEFINITION (Un-normalized, unamplified): ce Homo sapiens cDNA clone ce104h03
5', mRNA sequence.
VERSION CA389968
KEYWORDS EST. CA389968.1 GI:24720628
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Wistow,G., Bernstein,S.L., Wyatt,M.K., Paria,R.N., Behal,A.,
AUTHORS Touchman,J.W., Bouffard,G., Smith,D. and Peterson,K.
TITLE Expressed sequence tag analysis of human RPE/choroid for the
NEIBank Project: Over 6000 non-redundant transcripts, novel genes
and splice variants
JOURNAL Wistow G
MEDLINE Contact: Wistow G
PUBMED National Eye Institute
12107410 6/331, NIH, Bethesda, MD 20892-2740, USA
12103460 Tel: 301 402 3452
301 402 3452 Fax: 301 496 0078
Email: graeme@helix.nih.gov
Plate: 104 row: h column: 03
Seq primer: M13Rpl reverse primer (ABI).
FEATURES
source
1..639
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="ce104h03"
/tissue_type="RPE/choroid"
/dev_stage="Adult"
/lab_host="EMD10B"
/clone_lib="Human Retinal pigment epithelium/choroid cDNA
(un-normalized, unamplified): ce"
/note="Organ: Eye; Vector: pCMVSPORT; Two different donor
eyes (75-80 years old) yielded approximately 600 mg of
disected RPE/choroid tissue. This in turn yielded 340 ug
of total RNA and 7 ug of mRNA. A directionally cloned cDNA
library in the pCMVSPORT vector was constructed at life

Technologies (Rockville, MD; now part of Invitrogen Corp), essentially following the protocols of the SuperScript Plasmid System (Invitrogen Corp).
<http://www.invitrogen.com/> The library code designation was cs. For this library, cDNA inserts were cloned into the NotI/MluI sites of the vector. EST analysis was performed on the unamplified library at the NIH Intramural Sequencing Center (NISC)."

ORIGIN

Query Match 20.3%; Score 389.2; DB 6; Length 639;
Best Local Similarity 81.3%; Pred. No. 1.2e-90;
Matches 451; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

2 TGCAGGCAAGTACTATGACCTACCTACAAAGAGTACGAGTCCGCTGCTGCTC 61
Db TGCAGGCAAGTACTATGACCTACCTACAAAGAGTACGAGTCCGCTGCTGCTC 61
85 TGCAGGCAAGTACTATGACCTACCTACAAAGAGTACGAGTCCGCTGCTGCTC 144
Qy TGCAGGCAAGTACTATGACCTACCTACAAAGAGTACGAGTCCGCTGCTGCTC 121
Db TGCAGGCAAGTACTATGACCTACCTACAAAGAGTACGAGTCCGCTGCTGCTC 145
145 TGCAGGCAAGTACTATGACCTACCTACAAAGAGTACGAGTCCGCTGCTGCTC 204
Qy TGCAGGCAAGTACTATGACCTACCTACAAAGAGTACGAGTCCGCTGCTGCTC 181
Db TGCAGGCAAGTACTATGACCTACCTACAAAGAGTACGAGTCCGCTGCTGCTC 205
182 TGCAGGCAAGTACTATGACCTACCTACAAAGAGTACGAGTCCGCTGCTGCTC 241
241 TGCAGGCAAGTACTATGACCTACCTACAAAGAGTACGAGTCCGCTGCTGCTC 324
Qy TGCAGGCAAGTACTATGACCTACCTACAAAGAGTACGAGTCCGCTGCTGCTC 301
Db TGCAGGCAAGTACTATGACCTACCTACAAAGAGTACGAGTCCGCTGCTGCTC 325
325 TGCAGGCAAGTACTATGACCTACCTACAAAGAGTACGAGTCCGCTGCTGCTC 384
Qy TGCAGGCAAGTACTATGACCTACCTACAAAGAGTACGAGTCCGCTGCTGCTC 361
Db TGCAGGCAAGTACTATGACCTACCTACAAAGAGTACGAGTCCGCTGCTGCTC 385
385 TGCAGGCAAGTACTATGACCTACCTACAAAGAGTACGAGTCCGCTGCTGCTC 444
Qy TGCAGGCAAGTACTATGACCTACCTACAAAGAGTACGAGTCCGCTGCTGCTC 421
Db TGCAGGCAAGTACTATGACCTACCTACAAAGAGTACGAGTCCGCTGCTGCTC 445
445 TGCAGGCAAGTACTATGACCTACCTACAAAGAGTACGAGTCCGCTGCTGCTC 504
Qy TGCAGGCAAGTACTATGACCTACCTACAAAGAGTACGAGTCCGCTGCTGCTC 481
Db TGCAGGCAAGTACTATGACCTACCTACAAAGAGTACGAGTCCGCTGCTGCTC 505
505 TGCAGGCAAGTACTATGACCTACCTACAAAGAGTACGAGTCCGCTGCTGCTC 564
Qy TGCAGGCAAGTACTATGACCTACCTACAAAGAGTACGAGTCCGCTGCTGCTC 541
Db TGCAGGCAAGTACTATGACCTACCTACAAAGAGTACGAGTCCGCTGCTGCTC 565
565 TGCAGGCAAGTACTATGACCTACCTACAAAGAGTACGAGTCCGCTGCTGCTC 624
Qy TGCAGGCAAGTACTATGACCTACCTACAAAGAGTACGAGTCCGCTGCTGCTC 596
Db TGCAGGCAAGTACTATGACCTACCTACAAAGAGTACGAGTCCGCTGCTGCTC 625

RESULT 15
CV024947 491 bp mRNA linear EST 20-AUG-2004
LOCUS 2492 Full length cDNA from the Mammalian Gene Collection Homo
DEFINITION sapiens cDNA 5' similar to BC015220, mRNA sequence.
ACCESSION CV024947
VERSION CV024947.1 GI:51482835
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 491)
Rui, J.F., Hirozane-Kishikawa, T., Hao, T., Bertin, N., Li, S., Dricot, A., Li, N., Rosenberg, J., Lamesch, P., Vidalain, P.O., Clingingsmith, T.R., Hartley, J.L., Esposito, D., Cheo, D., Moore, T., Simmons, B., Segueria, R., Bosak, S., Doucette-Stamm, L., Le Peuch, C., Vandenhaute, J., Cusick, M.E., Albaladejo, J.S., Hill, D.E. and Vidal, M.

TITLE Human ORFome Version 1.1: a Platform for Reverse Proteomics
JOURNAL Genome Res. (2004) In press
COMMENT Contact: Vidal M
Marc Vidal Laboratory
Dana Farber Cancer Institute
1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA
Tel: 617 632 5180
Fax: 617 632 5739
Email: Marc.Vidal@dfci.harvard.edu

ORF Sequence Tag (OST) of Gateway construct. Each cloned ORF results from a PCR reaction using an MGC full-length cDNA as template DNA and ORF specific primers
PCR Primers
FORWARD: ATGTTGAGAACTGACTCTGAT
BACKWARD: TAGGAATGCTTCATCCCTG
Insert Length: 491 Std Error: 39.00
Plate: 11066 Row: 05 Column: G
Seq primer: ACTGCGCGTGTGTTACAGAGTGTGATGAGGAAAC
High quality sequence start: 94
High quality sequence stop: 490
POLYA=No.

FEATURES

SOURCE

Location/Qualifiers
1..491
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="Full length cDNA from the Mammalian Gene Collection"
/note="Vector: mixed; The ORFs were PCR amplified from the MGC (Mammalian Gene Collection) as of April 2004 and cloned by recombinational Gateway cloning into pDONR223 Donor Vector. Reference: MGC (Mammalian Gene Collection) Program Team. Generation and Initial Analysis of more than 15,000 Full-length Human and Mouse cDNA Sequences. PNAS, 2002, 99(26), 16899-16903"

ORIGIN

Query Match 19.9%; Score 381.2; DB 7; Length 491;
Best Local Similarity 86.1%; Pred. No. 1.5e-88;
Matches 422; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

192 TGTGTTGAGAAAGTGGCTCTGTAAGTACGAGTACATTCAGTATCCCTATTCCTTCG 251
Db TGTGTTGAGAAAGTGGCTCTGTAAGTACGAGTACATTCAGTATCCCTATTCCTTCG 251
2 TGTGTTGAGAAAGTGGCTCTGTAAGTACGAGTACATTCAGTATCCCTATTCCTTCG 61
Qy TGTGTTGAGAAAGTGGCTCTGTAAGTACGAGTACATTCAGTATCCCTATTCCTTCG 311
Db TGTGTTGAGAAAGTGGCTCTGTAAGTACGAGTACATTCAGTATCCCTATTCCTTCG 62
62 TGTGTTGAGAAAGTGGCTCTGTAAGTACGAGTACATTCAGTATCCCTATTCCTTCG 121
Qy TGTGTTGAGAAAGTGGCTCTGTAAGTACGAGTACATTCAGTATCCCTATTCCTTCG 371
Db TGTGTTGAGAAAGTGGCTCTGTAAGTACGAGTACATTCAGTATCCCTATTCCTTCG 122
122 TGTGTTGAGAAAGTGGCTCTGTAAGTACGAGTACATTCAGTATCCCTATTCCTTCG 181
Qy TGTGTTGAGAAAGTGGCTCTGTAAGTACGAGTACATTCAGTATCCCTATTCCTTCG 431
Db TGTGTTGAGAAAGTGGCTCTGTAAGTACGAGTACATTCAGTATCCCTATTCCTTCG 182
182 TGTGTTGAGAAAGTGGCTCTGTAAGTACGAGTACATTCAGTATCCCTATTCCTTCG 241
Qy TGTGTTGAGAAAGTGGCTCTGTAAGTACGAGTACATTCAGTATCCCTATTCCTTCG 491
Db TGTGTTGAGAAAGTGGCTCTGTAAGTACGAGTACATTCAGTATCCCTATTCCTTCG 242
242 TGTGTTGAGAAAGTGGCTCTGTAAGTACGAGTACATTCAGTATCCCTATTCCTTCG 301
Qy TGTGTTGAGAAAGTGGCTCTGTAAGTACGAGTACATTCAGTATCCCTATTCCTTCG 551
Db TGTGTTGAGAAAGTGGCTCTGTAAGTACGAGTACATTCAGTATCCCTATTCCTTCG 302
302 TGTGTTGAGAAAGTGGCTCTGTAAGTACGAGTACATTCAGTATCCCTATTCCTTCG 361
Qy TGTGTTGAGAAAGTGGCTCTGTAAGTACGAGTACATTCAGTATCCCTATTCCTTCG 611
Db TGTGTTGAGAAAGTGGCTCTGTAAGTACGAGTACATTCAGTATCCCTATTCCTTCG 362
362 TGTGTTGAGAAAGTGGCTCTGTAAGTACGAGTACATTCAGTATCCCTATTCCTTCG 421
Qy TGTGTTGAGAAAGTGGCTCTGTAAGTACGAGTACATTCAGTATCCCTATTCCTTCG 671
Db TGTGTTGAGAAAGTGGCTCTGTAAGTACGAGTACATTCAGTATCCCTATTCCTTCG 422

Qy 672 GTGGACACT 681
|||
Db 482 GTGGACACT 491

Search completed: March 26, 2005, 13:32:12
Job time : 6165.57 sec8

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2005, 21:57:50 ; Search time 993.266 Seconds
(without alignments)
11419.125 Million cell updates/sec

Title: US-09-622-964A-28

Perfect score: 1916
Sequence: 1 gtcgcaagcgcgtactca.....aaaaaaaaaaaaaaaaaa 1916

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: _geneseq1980s:*
2: _geneseq1990s:*
3: _geneseq2000s:*
4: _geneseq2001s:*
5: _geneseq2001bs:*
6: _geneseq2002as:*
7: _geneseq2002bs:*
8: _geneseq2003as:*
9: _geneseq2003bs:*
10: _geneseq2003cs:*
11: _geneseq2003ds:*
12: _geneseq2004as:*
13: _geneseq2004bs:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1916	100.0	1916	2	AAZ21229 Mouse CG1
2	871.2	45.5	2229	2	AAZ21227 Human CG1
3	867	45.3	1758	8	ABZ80972 Human bes
4	658.2	34.4	2404	12	ADQ84435 Human tum
5	658.2	34.4	2404	13	ADQ83262 Human tum
6	658.2	34.4	2429	2	AAZ21228 Human CG1
7	519.4	27.1	2112	6	ABQ61177 FLJ20132
8	519.4	27.1	2137	8	ABZ80973 Human bes
9	519	27.1	1530	8	ABZ80997 Human bes
10	458.2	23.9	2500	11	ADM02614 Human bes
11	423	22.1	2028	8	ABZ80974 Human bes
12	401.2	20.9	1948	9	ACC59900 Human REM
13	399.2	20.8	1422	8	ABZ80975 Human bes
14	296.2	15.5	2861	4	ABL10793 Drosophila
15	260.4	13.6	1526	5	AAZ72796 DNA encod
16	229	12.0	1608	4	ABL12609 Drosophila
17	227.6	11.9	1263	2	AAV99722 Human adu
18	215.4	11.2	10760	4	ABL10792 Drosophila
19	211.2	11.0	1592	10	ADC29982 Human nov
20	194.4	10.1	620	5	ADL45614 Human ova

21	194.4	10.1	1345	4	ABL12637 Drosophila
22	190.8	10.0	305	5	ABAI1094 Human ner
23	188.4	9.8	7108	5	ABAI4556 Human ner
24	188.4	9.8	16125	2	AAZ21226 Human CG1
25	188.4	9.8	16650	5	ABAI4559 Human ner
26	188.4	9.8	18530	5	ABAI4557 Human ner
27	188.4	9.8	18537	5	ABAI4558 Human ner
28	187.8	9.8	1350	4	AAZ76848 Human sec
29	187.2	9.8	2196	4	ABL16195 Drosophila
30	181.2	9.5	1292	11	ACN89920 Breast ca
31	166.6	8.7	3592	4	ABL12636 Drosophila
32	163	8.5	1238	9	ADA44960 Human pol
33	129.6	6.8	3785	12	ACH74388 Human gen
34	122.2	6.4	539	12	ACH88088 Human gen
35	119.8	6.3	235	12	ADQ24621 Human sof
36	118	6.2	2914	12	ADQ25463 Human sof
37	118	6.2	853	11	ACN90196 Breast ca
38	114.4	6.0	1815	10	ADQ24756 Human sof
39	113.8	5.9	1954	12	ADQ24756 Human sof
40	113.4	5.9	1330	10	ADC31917 Human nov
41	105.6	5.5	2104	5	AAZ72795 DNA encod
42	105.6	5.5	3592	4	ABL12636 Drosophila
43	95.4	5.0	527	4	ABL17589 Human bre
44	95	5.0	527	4	ABL17589 Human bre
45	94.4	4.9	392	4	ABL11210 Human bre

ALIGNMENTS

RESULT 1	
ID	AAZ21229 standard; CDNA; 1916 BP.
XX	AAZ21229;
AC	22-NOV-1999 (first entry)
DT	Mouse CG1CE CDNA sequence.
XX	
XX	CG1CE; Best's macular dystrophy; mutation; diagnosis; detection; BMD;
KW	age-related macular dystrophy; se.
XX	
OS	Mus sp.
XX	
FT	Location/Qualifiers
FT	11..166
FT	/*tag= a
FT	/product= "CG1CE protein"
XX	
PN	W09943695-A1.
XX	
PD	02-SEP-1999.
XX	
PF	22-FEB-1999; 99WO-US003790.
XX	
PR	25-FEB-1998; 98US-0075941P.
PR	18-DEC-1998; 98US-0112926P.
XX	
PA	(MERI) MERCK & CO INC.
PA	(UYUP-) UNIV UPPSALA.
XX	
PI	Petrushin K, Caskey CT, Metzger M, Madeline C;
XX	
DR	WP1; 1999-540560/45.
DR	P-PSDB; AAY29955.
XX	
PT	Human and mouse polynucleotides encoding CG1CE polypeptides.
XX	
XX	Claim 2; Fig 8; 67pp; English.
XX	
CC	The present sequence represents the mouse CG1CE CDNA sequence, which when mutated is responsible for Best's macular dystrophy (BMD).

CC Polynucleotides encoding CGICE are useful for diagnosing whether a
CC patient carries a mutation in the CGICE gene. Normal and mutated CGICE
CC proteins are useful for identifying activators and/or inhibitors of these
CC proteins, in order to treat BMD. The CGICE gene offers a simpler and
CC cheaper method of diagnosing BMD without the need for the presence of the
CC patient. The gene may also be useful to discovering the genetic cause of
CC age-related macular dystrophy
XX

Sequence 1916 BP; 514 A; 499 C; 457 G; 446 T; 0 U; 0 Other;

Query Match 100.0%; Score 1916; DB 2; Length 1916;
Best Local Similarity 100.0%; Pred.No. 0;
Matches 1916; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGCCAAGCATGATCATCTATCAACAACAAGTACCAATGCGCGCTCGTTCCT 60
DB 1 GTGCCAAGCATGATCATCTATCAACAACAAGTACCAATGCGCGCTCGTTCCT 60
QY 61 CTGCTCCCTCTCTGCTGCTGCGAGGAGCATCTACAAGCTGTGTATGAGAAATTCCT 120
DB 61 CTGCTCCCTCTCTGCTGCTGCGAGGAGCATCTACAAGCTGTGTATGAGAAATTCCT 120
QY 121 TGTCTCAATATCCCTCTACTATTCATCCGTGAGACTCTACAGATGTTCTCTGAGTGA 180
DB 121 TGTCTCAATATCCCTCTACTATTCATCCGTGAGACTCTACAGATGTTCTCTGAGTGA 180
QY 181 TCAGCAGCTGTGTTGAGAAAGCTGCTCTGACTGCGACAGCTACATTCAGCTCATCC 240
DB 181 TCAGCAGCTGTGTTGAGAAAGCTGCTCTGACTGCGACAGCTACATTCAGCTCATCC 240
QY 241 TATATCTCTCGTTCGTGGTTTCTATGTTACATTTGTGTGAGCCGCTGTGAGCCAGTA 300
DB 241 TATATCTCTCGTTCGTGGTTTCTATGTTACATTTGTGTGAGCCGCTGTGAGCCAGTA 300
QY 301 CGAGAAGCTTGCCTGCGGCCGACCGGCTCATATCCAGAGTCTTACCTTCGTGAGAGGCA 360
DB 301 CGAGAAGCTTGCCTGCGGCCGACCGGCTCATATCCAGAGTCTTACCTTCGTGAGAGGCA 360
QY 361 GGATGAGAGAGGCGCTTGTGTGCGGCGACCGCTCATCCGCTACGCGCATCTGGGCGCAAGT 420
DB 361 GGATGAGAGAGGCGCTTGTGTGCGGCGACCGCTCATCCGCTACGCGCATCTGGGCGCAAGT 420
QY 421 GCTCATCTCTGCGAGCATCAGCACTCGGTCTACAGCGCTTCCACTTTCACCACT 480
DB 421 GCTCATCTCTGCGAGCATCAGCACTCGGTCTACAGCGCTTCCACTTTCACCACT 480
QY 481 GGTGTACAGAGTTTATGACCAATGGGGAACATTAAGAGTTGAGAAAGTGGGCTTACC 540
DB 481 GGTGTACAGAGTTTATGACCAATGGGGAACATTAAGAGTTGAGAAAGTGGGCTTACC 540
QY 541 ACACAACACATTTGGGTCCTGCGGTGTGTTGCCAACTTGTCAATGAAGGCTTACT 600
DB 541 ACACAACACATTTGGGTCCTGCGGTGTGTTGCCAACTTGTCAATGAAGGCTTACT 600
QY 601 TGGAGTCTGATCCGGGACACCGTCTGCTCAAGAGCTGTGAATGAAGTGTACTTT 660
DB 601 TGGAGTCTGATCCGGGACACCGTCTGCTCAAGAGCTGTGAATGAAGTGTACTTT 660
QY 661 GCGTACTCAGTGGAGCAGCTGTATGCTCAAGCTGTGAATGAAGTGTACTTT 720
DB 661 GCGTACTCAGTGGAGCAGCTGTATGCTCAAGCTGTGAATGAAGTGTACTTT 720
QY 721 ACAGTGTGTGACAGTGTGAGTATACAGTTTCTTTCATGCTGTGAATGGAGGAGGTT 780
DB 721 ACAGTGTGTGACAGTGTGAGTATACAGTTTCTTTCATGCTGTGAATGGAGGAGGTT 780
QY 781 TCTGAACCCCAAAACAAGACTACCAAGGCGCAGTGAATCTGTGTGTGCTGTTCAC 840
DB 781 TCTGAACCCCAAAACAAGACTACCAAGGCGCAGTGAATCTGTGTGTGCTGTTCAC 840
QY 841 AATCTGCAATTTCTTATCTACATGGGCTGAGTAAGGTGAGAAAGCTCATCAACCC 900
DB 841 AATCTGCAATTTCTTATCTACATGGGCTGAGTAAGGTGAGAAAGCTCATCAACCC 900

QY 901 CTTGCGGAGAGACGATGATATTTTGAAGCTAATCTGATCATTTGACAGAAACCTTGAGGT 960
DB 901 CTTGCGGAGAGACGATGATATTTTGAAGCTAATCTGATCATTTGACAGAAACCTTGAGGT 960
QY 961 GTCCCTGTGTCCGTGATGAGATGACACGAATCTGCTCCATGAGAAAGTGCATGTA 1020
DB 961 GTCCCTGTGTCCGTGATGAGATGACACGAATCTGCTCCATGAGAAAGTGCATGTA 1020
QY 1021 CTGAAAGAGGACGCGCTTACGCGCTTACACAGCTGTCTTCTGCGAGTCTGCGCGCA 1080
DB 1021 CTGAAAGAGGACGCGCTTACGCGCTTACACAGCTGTCTTCTGCGAGTCTGCGCGCA 1080
QY 1081 TTCCCTATGAGGCTCCACTTTCACATCAGCTTAAAGAAAGAACTTAAGCTTTGCTC 1140
DB 1081 TTCCCTATGAGGCTCCACTTTCACATCAGCTTAAAGAAAGAACTTAAGCTTTGCTC 1140
QY 1141 AAAAGAGAGGCTGACACGAGTAAAGAAAGAGTGGTATAGACACCATAGGCTGCTT 1200
DB 1141 AAAAGAGAGGCTGACACGAGTAAAGAAAGAGTGGTATAGACACCATAGGCTGCTT 1200
QY 1201 CTTAGGACTGCAACCCCAAACTTACATCTTCCCTTGAAGACTTAAAGCCAACTATT 1260
DB 1201 CTTAGGACTGCAACCCCAAACTTACATCTTCCCTTGAAGACTTAAAGCCAACTATT 1260
QY 1261 GTGTTTAAAGAACCCCTCTCTGAAAGGCGAGTGAAGATGGCAACCAAGAAACAGAA 1320
DB 1261 GTGTTTAAAGAACCCCTCTCTGAAAGGCGAGTGAAGATGGCAACCAAGAAACAGAA 1320
QY 1321 AGATGTCTGAAATTTAAAGGCTGTGACTTTTGAATGTGTTCAAGGTTTAAAGAGAG 1380
DB 1321 AGATGTCTGAAATTTAAAGGCTGTGACTTTTGAATGTGTTCAAGGTTTAAAGAGAG 1380
QY 1381 AGGCTCCATTTGTGGCCCAAGGACCCAGACGCCCTTACTGAGCAGTCAAGCACCCTC 1440
DB 1381 AGGCTCCATTTGTGGCCCAAGGACCCAGACGCCCTTACTGAGCAGTCAAGCACCCTC 1440
QY 1441 CAGTTGACACAGGATGAGGCTTCAACAGATTACAAAGAAATCTGTCACTGAAAAA 1500
DB 1441 CAGTTGACACAGGATGAGGCTTCAACAGATTACAAAGAAATCTGTCACTGAAAAA 1500
QY 1501 GAAAACTGTGAGTTTAACTTGAACATTCAGAGAGCCCAAGAACATCTTCAACAGCG 1560
DB 1501 GAAAACTGTGAGTTTAACTTGAACATTCAGAGAGCCCAAGAACATCTTCAACAGCG 1560
QY 1561 CCGTTTGAACAGATGTCAACCAATTAACAGGCTTAAATGAAGAGCAGTCAAGTCTTA 1620
DB 1561 CCGTTTGAACAGATGTCAACCAATTAACAGGCTTAAATGAAGAGCAGTCAAGTCTTA 1620
QY 1621 TCCCTACAGGATGAGCTGCGACCAAACTGTTCTTATGAGTGTGCTCAGAGCTTG 1680
DB 1621 TCCCTACAGGATGAGCTGCGACCAAACTGTTCTTATGAGTGTGCTCAGAGCTTG 1680
QY 1681 GCCCTGACTTGAAGAGTCCAGAGGCGCATGACCCAGTCAAGGCAACAGAGCG 1740
DB 1681 GCCCTGACTTGAAGAGTCCAGAGGCGCATGACCCAGTCAAGGCAACAGAGCG 1740
QY 1741 ACACCCAGAGAGTGTGTTCCAGACAGCTAGAGTAACTCAGAAACCAAGATCTTAA 1800
DB 1741 ACACCCAGAGAGTGTGTTCCAGACAGCTAGAGTAACTCAGAAACCAAGATCTTAA 1800
QY 1801 TAGTCTGCTGAAAAACCTGTATTTTTCATGCTTTCCTCAAACTAAGAGTTTAAATA 1860
DB 1801 TAGTCTGCTGAAAAACCTGTATTTTTCATGCTTTCCTCAAACTAAGAGTTTAAATA 1860
QY 1861 CCGTAATATTTCTTTTGTGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1916
DB 1861 CCGTAATATTTCTTTTGTGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1916

RESULT 2
AAZ21227
ID AAZ21227 standard; cDNA; 2229 BP.

XX	AA221227;
AC	
XX	
DT	22-NOV-1999 (first entry)
XX	
DE	Human CGICE short form cDNA sequence.
XX	
KW	CGICE; Beet's macular dystrophy; mutation; diagnosis; detection; BMD;
XX	
KW	age-related macular dystrophy; ss.
XX	
OS	Homo sapiens.
XX	
FT	Key
FT	Location/Qualifiers
FT	CDS
FT	105..1862
FT	/*tag= a
FT	/product= "CGICE long form protein"
FT	/trans_except= (pos:465..467,aa:Ser)
XX	
PN	W09943695-A1.
XX	
PD	02-SEP-1999.
XX	
PP	22-FEB-1999; 99WO-US003790.
XX	
PR	25-FEB-1998; 98US-0075941P.
PR	18-DEC-1998; 98US-0112926P.
XX	
PA	(MERI) MERCK & CO INC.
PA	(UYUP-) UNIV UPPSALA.
XX	
PI	Petrukhin K, Caskey CT, Metzker M, Wadelius C,
XX	
DR	WPI; 1999-540560/45.
DR	P-PSDB; AAY29953.
XX	
PT	Human and mouse polynucleotides encoding CGICE polypeptides.
XX	
XS	Claim 2; Fig 2; 67pp; English.
XX	
CC	The present sequence represents the human CGICE cDNA sequence, which when
CC	mutated is responsible for Beet's macular dystrophy (BMD).
CC	Polynucleotides encoding CGICE are useful for diagnosing whether a
CC	patient carries a mutation in the CGICE gene. Normal and mutated CGICE
CC	proteins are useful for identifying activators and/or inhibitors of these
CC	proteins, in order to treat BMD. The CGICE gene offers a simpler and
CC	cheaper method of diagnosing BMD without the need for the presence of the
CC	patient. The gene may also be useful to discovering the genetic cause of
CC	age-related macular dystrophy
XX	
Q0	Sequence 2229 BP; 575 A; 646 C; 532 G; 476 T; 0 U; 0 Other;

Query Match	45.5%	Score 871.2	DB 2	Length 2229
Beet Local Similarity	77.1%	Pred. No. 1.2e-210		
Matches 1116	Conservative	0	Mismatches 293	Indels 39
			Gaps	3
Oy	2	TGCCAAGCATGATCATCATCAACAACAAAGTAGCAATAGCCCGCTCGATTGCTTC	61	
Db	96	TGCCGCGCATATCCATCATCTACACACAGCCAAAGTGGCTAATGCCGCTTAGGCTCTTC	155	
Oy	62	TGTCCTCTCTCTGTGTGGGAGGAGCATCTACAAAGCTCTGTATGAGAAATTCCTT	121	
Db	156	TCCCGCTGTGTGTGTGGGAGGAGCATCTACAAAGCTGTATATGGCGAGTCTTAA	215	
Oy	122	GTCTTCATATTCCTCTACTATTCATTCCTGTGAGACTTACAGAAATGGTTCTTCGATGAT	181	
Db	216	ATCTTCTCGCTCTGTCTACTACATCATCCGCTTATTTAATGAGCTGGCCCTCAGGAAGAA	275	
Oy	182	CAGCAGCTGTTGTTGAGAAAGCTGGCTGTACTGTGAGCAGGTACATTCAGCTCATCCCT	241	
Db	276	CAACAGCTGATGTTTGAGAAAGCTGACTCTGTATTTGGAGACAGCTACATCCAGCTCATCCCC	335	
Oy	242	ATATCCTTCGTTCTGGGGTTCTATATGTTACATTTGGTGTGAGCCGCTGTGTGAGACCAGTAC	301	

Db	336	ATTTCCTTGTGCTGGGGCTTTCAAGTGAAGCTGGATCGTGAACCCGCTGGTGGAAACAGTAC	395
QY	302	GAGAACTTGGCGTGCCCGGACCGGCTTCAATATCCAGAGTGTCTAGCTTCTGTGAAGGCAAG	361
Db	336	GAGAACTTGGCGTGCCCGGACCGGCTTCAATATCCAGAGTGTCTAGCTTCTGTGAAGGCAAG	455
QY	352	GATAGGGAAGGCGTTTGTCTGGGGGAGAGGCTCATCCGCTAGGCAACTCTGGGCGCAAGT	421
Db	456	GACGAGCAAGGCGGCTCTGGGGGAGAGGCTCATCCGCTAGGCAACTCTGGGCGCAAGT	515
QY	422	CTCATCTGCGAGCATGACGACCTCGGCTTCAAGAGCTTTCCGACTTTCACACACTG	481
Db	516	CTCATCTGCGAGGCTGACGACCGGCAAGTCAAGAGCTTTCCGAGCGCCCGACGACCTG	575
QY	482	GTGCTAGCAGGTTTTATGACCCATGGGGAACATAAAGAGTTGTCAGAAATTGGGCTTACCA	541
Db	576	GTGCAAGCAGGCTTATGATCTCGGACAGAACAAAGAGTTGAGAAATCGAGCCCTACCA	635
QY	542	CACAAACAATTCTGGGTCCTCGGCTGTGGTGTGGCAACTTGTCAATGAAGGCTTACTT	601
Db	636	CACAACAATGTTCTGGGTCCTCGGCTGTGGTGTGGCAACTTGTCAATGAAGGCGTGGCTT	695
QY	602	GGAGGTGCAATCCGGGACACCGTCTGTGCTCCAGAGCTGATGAATGAGGTGTGTAATTG	661
Db	686	GGAGGTGCAATCCGGGACCTTACTGTGCTCCAGAGCTGCTGAACGATGAACACTTG	755
QY	662	CGTACTCAGTGTGACAGCTGTATGCTCAACGATGGAATAGTATCCACTTGGTGTATACACA	721
Db	756	CGTACTCAGTGTGACACTGTATGCTCAACGATGGAATAGTATCCACTTGGTGTATACACA	815
QY	722	CAGTGTGTGACAGTGGCAGTATACAGCTTTTTCCTTGCATGCTTGATCGGAGGCGAGTTT	781
Db	816	CAGTGTGTGACGTGGCGGCTGTACAGCTTCTTGCATGCTTGATGGGCGGCGAGTTT	875
QY	782	CTGAACCCAAACAAAGACATACCCAGGACATGAGATCTGAGTGTGCTGTCTTACACA	841
Db	876	CTGAACCCAGCCAAAGGCTTACCTTGGCCATGAGCTGAGACTTGTGTGCTTCTTACAG	935
QY	842	ATCTGTCAATTTTATTTCTACATGGGCTGGCTGAAGGTGGCAACAGCTCATCAACCC	901
Db	936	TTCTGTGAGTTTCTTCTATGTGTGGCTGGCTGAAGGTGGCAAGGCTCATCAACCC	995
QY	902	TTTGGGGAGACGATGATGATTTTGAAGTAACTGAGTCAATTGACAGAAACTGCAAGT	961
Db	996	TTTGGAGAGATGATGATTTTGAAGCAAACTGATTTGTGCACAGAAATTTGCAAGT	105
QY	962	TCCCTGTGTCCGTGGATGGGATGGAACCAAGACTTGGCTCCATGGAAAGTACATGTAAC	102
Db	1056	TCCCTGTGTGTGGATGAGATGGAACCAAGACTTGGCTCCATGGAAAGTACATGTAAC	111
QY	1022	TGGAAAGAGGCAAGGCTCTAGCGGCGCTTACACAGCTGTCTTGCCAGGTCTGAGCGGCAAT	108
Db	1116	TGGAAATTAAGCCCGAGGCTCAGGCGCTTACACAGCTGTCTTGCCAGGTCTGAGCGGCAAT	117
QY	1082	TCCTTCATGGGCTCCACCTTCAACATCAGCTTAAGAAAGAAAGAACTTGAAGCTTTGGTCA	114
Db	1176	TCCTTCATGGGCTCCACCTTCAACATCAGCTTAAGAAAGAAAGAAAGTGTTCAGGCC	123
QY	1142	AAAGAGAGGCTGACACGATTAAGAAAGAGAGTGGCTATGACAGCACTTATGGCTGCTC	120
Db	1236	AATCAGGAGGAGAGAGAGATGCTCAGCTG-----GCATCTATGGCGGCTTTC	128
QY	1202	TTAGGACATGCAACCCAAATACTAACATCTTCCCTTGAAGAATTAAAGACCAATATG	126
Db	1284	CTTAGGCTGTGAGTCCCATGATCACAATCTCTCCAGGCGCAATCTCAAGACCAATCTAGT	134
QY	1262	TGTTCTTAAGA-----ACCCCTCTCTGAAGGCGAGTGAAGATGCCAACCAAGAAAAC	131
Db	1344	TGGCCCAAGAGGAATCCCTTCTTCAAGGAGGCTGCGCCAAAACCAAGAGCAGCAAA	140
QY	1316	CAGAAAG-----ATGTCTGGAATTTTAAGGCTTGAATCTTCTTG	135
Db	1404	CAGAACTTGAAGGCGCAGAGAAACAACAAGGCTTGAAGCTTAAAGCTGTGACGCTTTC	146

QY 1355 AATGTGTTCCAGGTAAAGAGAGAGGCTCCATTGTGCCCCACAGGACCAGC 1414
DB 1464 AAGCTGCCCCACTGTATTCAGAGGCGACTACTACGTGCCCCACAGAGCCCTTCAGC 1523
QY 1415 CACCTTAC 1422
DB 1524 CCCACTCC 1531

RESULT 3
AB280972
ID AB280972 standard; DNA; 1758 BP.
XX
AC AB280972;
XX
DT 15-OCT-2003 (first entry)
XX
DE Human bestrophin (vitelliform macular dystrophin 2) gene.
XX
KW Bestrophin; vitelliform macular dystrophin 2; VMD2; ds; gene; vaccine;
KW antiobesity; antidiabetic; immunomodulator; hypotensive; cardiac;
KW antileptic; osteopathic; antiinflammatory; cyostatic; obesity;
KW energy homeostasis; metabolism; triglyceride; body-weight regulation;
KW eating disorder; cachexia; diabetes mellitus; hypertension; gallstones;
KW coronary heart disease; hypercholesterolemia; osteoarthritis; cancer;
KW Sleep apnea; chromosome 11.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1758
FT FT /tag= a
FT FT /product= "VMD2"

XX WO200303922-A2.
XX PD 17-APR-2003.
XX
XX 09-OCT-2002; 2002WO-BP011321.
XX PR 09-OCT-2001; 2001EP-00124059.
XX
XX (DEVE-) DEVELOPENTWICKLUNGSHIOLOGISCHE FORSCH.
XX
XX Steernagel A, Broenner G, Fritsch R, Eulenberg K, Ciosek T;
XX WPI: 2003-393411/37.
XX P-PSDB; ABR58055.
XX
XX New pharmaceutical compositions comprising a Bestrophin gene, polypeptide
XX or nucleic acid, for treating, alleviating and/or preventing metabolic
XX diseases, e.g. obesity, cachexia, diabetes mellitus, hypertension, or
XX gallstones.
XX
XX Claim 2; Fig 4a; 85pp; English.
XX
XX This sequence represents the coding region for a member of the human
XX bestrophin gene family designated vitelliform macular dystrophin 2
XX (VMD2). The dystrophin gene family are involved in energy homeostasis and
XX metabolism of triglycerides. The sequence can be used for the manufacture
XX and/or preventing disorders including metabolic diseases such as obesity
XX and other body-weight regulation and related disorders such as eating
XX disorder, cachexia, diabetes mellitus, hypertension, coronary heart
XX disease, hypercholesterolemia, osteoarthritis, gallstones, cancers of the
XX reproductive organs, and sleep apnea. The gene is found on human
XX chromosome 11
XX
XX Sequence 1758 BP; 421 A; 515 C; 448 G; 374 T; 0 U; 0 Other;
XX
XX Query Match 45.3%; Score 867; DB 8; Length 1758;
XX Best Local Similarity 77.1%; Pred. No. 1.3e-209;

Matches 1110; Conservative 0; Mismatches 290; Indels 39; Gaps 3;
QY 11 ATGACTACCTTACACAAACAAAGTAGCCAAATGCGGCTCGGTTCTGTCCCTC 70
DB 1 ATGACATCATCTTACACAGCCAAAGTGTCTAATGCCCTTAGGCTCTTCTCCCGCTG 60
QY 71 CTCCTGTGCTGCGAGGCGAGCATCTACAGCTGTGTATGAGAAATTCCTTGTCTTATA 130
DB 61 CTGCTGTGCTGCGGCGAGCATCTACAGCTGTATATGCGAGTTCTTAATCTTCTG 120
QY 131 TTCCTACTATTCATCCCGTGAAGCTTACAGAAATGTTCTCTGAGTATCGAGCTG 190
DB 121 CTCCTACTATCATATCCCTTATTTATTAAGCTGCGCTCAGAGAAACAAAGCTG 180
QY 191 TTGTTGAGAACTGCTGTGTATCTGAGCAGCTACATTCAGCTCATCCCTATATCTTC 250
DB 181 ATGTTTGAAGAACTGACTGTGTATGTGAGACGTACATCCAGCTCATCCCATTTCTTC 240
QY 251 GTTCTGAGGTTTCTATATGTATACATTTGTGTGAGCGCTGTGTGAGCCAGTAAGAACTTG 310
DB 241 GTGCTGAGCTTCTTACGTGAGCTGTGTGTGACCCGCTGTGTGAAACAGTAACAGAACTTG 300
QY 311 CCGTGCCGAGCGCTCATGATCCAGGTGTCTAGCTTGTGTGAGGCGAAGATAGAGAA 370
DB 301 CCGTGCCGAGCGCTCATGATCCAGGTGTGTGTGAGGCGAAGATAGAGAACTTG 360
QY 371 GCGCGTTTGTGCGGCGAGCTCATGCTGCTACGCGCATCTGCGGCGCAAGTGTCTCATCTG 430
DB 361 GCGCGGCTGTGCGGCGAGCTCATGCTGCTACGCGCATCTGCGGCGCAAGTGTCTCATCTG 420
QY 431 CGCAGATACAGACCTCGGTCTTAAAGCGCTTTCCACCTTTCACACCTGTGTCTAGCA 490
DB 421 CGCAGCTACAGACCTCGGTCTTAAAGCGCTTTCCACCTTTCACACCTGTGTCTAGCA 480
QY 491 GGTTTTATAGACCATGAGGAGATTAAGCATTAAGCTTGAAGTTGAGCTTACACAAACA 550
DB 481 GGTTTTATAGACCATGAGGAGATTAAGCATTAAGCTTGAAGTTGAGCTTACACAAACA 540
QY 551 TTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 610
DB 541 TTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
QY 611 ATCCGCGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 670
DB 601 ATCCGCGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
QY 671 TGTGACACCTGTATGCTTACGATGTATGATATCCATTTGATGATACACAGGTGTG 730
DB 661 TGTGACACCTGTATGCTTACGATGTATGATATCCATTTGATGATACACAGGTGTG 720
QY 731 ACAGTGACAGTATACAGCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 790
DB 721 ACTGTGAGGCTGTATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
QY 791 AACAAAGACTACCCAGGCGATGAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 850
DB 781 GCCAAGGCTTACCTGCGCATGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
QY 851 TTCTTATCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 910
DB 841 TTCTTCTTCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
QY 911 GACGATGATGATTTTGAAGCTTACGATGATGATGATGATGATGATGATGATGATGATGAT 970
DB 901 GATGATGATGATTTTGAAGCTTACGATGATGATGATGATGATGATGATGATGATGATGAT 960
QY 971 TCCGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1030
DB 961 GCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
QY 1031 GCAGCGCTGAGCGGCTTACAGAGCTGCTTCTGCAAGCTGCTGCGGAGCTTCTTCTGATG 1090
DB 1021 CCGAGGCGACAGCGGCTTACAGAGCTGCTTCTGCAAGCTTCTGCGAGGCTTCTTATG 1080

QY 1091 GGCTCACTTCAACATCAGCCTTAAGAAAGAAAGACTTGGCTCAAAAGAGAG 1150
DB 1081 GGCTCACTTCAACATCAGCCTTAAGAAAGAAAGAGACTTGGCTCAAAAGAGAG 1140
QY 1151 GCTGACACGATTAAGAAAGAGAGCTTAAGAAAGAGACTTGGCTCAAAAGAGAG 1210
DB 1141 G-----ACGAGAGAGATGCTCAGCCTGCAATCAATGAGCTTCTTGAAGCTG 1188
QY 1211 CAACCAAAATCTCATCTTCCCTTGAAGACTTAAGAAAGAAAGACTTGGCTCAAAAGAGAG 1270
DB 1189 CAGTCCATGATCAACATCTCTCCAGGCAAACTCAAGAGCAAACTGAGGCTCAAG 1248
QY 1271 A-----ACCCCTCTCTGGAAGGCTTAAGAGTCCCAACCAAGAAAGAGAG-- 1322
DB 1249 AGGGAATCTCTCTCAAGAGGCTGCTCAAAAGCAAAAGGCAAGCAAGAGCTT 1308
QY 1323 -----ATGCTGAAATTTAAGGCTTGAAGCTTCTTGAATGTTT 1363
DB 1309 AGGGGCGAGAAAGCAACAAAGGCTGAAAGCTTAAGAGCTGAGACGCTTCAAGTCTGC 1368
QY 1364 CCAAGCTTTAAGAGAGAGCTTCCATGTCGCGCAAGGCAAGCAAGAGCAAGCTT 1422
DB 1369 CCACTGATACAGAGGCTCACTCAAGTGCCTCAGACAGACCTTCAAGCTTCC 1427

RESULT 4
AD084435
ID AD084435 standard; cDNA; 2404 BP.
AC AD084435;
DT 07-OCT-2004 (first entry)
XX
DE Human tumour-associated antigenic target (TAT) cDNA sequence #1249.
XX
XX human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;
KM cancer; cell proliferative disorder; gene; ss.
XX
OS Homo sapiens.
PN W02004060270-A2.
PD 22-JUL-2004.
XX
PF 15-OCT-2003; 2003WO-US029126.
XX
PR 18-OCT-2002; 2002US-0418988P.
XX
PA (GENE) GENENTECH INC.
PA (WUTD/) WU T D.
PA (ZHOU/) ZHOU Y.
PI Wu TD, Zhou Y;
PI WPI; 2004-534306/51.
XX
PT New nucleic acid molecule and encoded polypeptide, for diagnosing,
PT preventing or treating cell proliferative disorders such as cancer.
PS Claim 1; SEQ ID NO 1249; 5504bp; English.
XX
XX The present invention describes an isolated tumour-associated antigenic
CC target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide
CC sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of
CC (a); (c) the complement of (a) or (b); (d) a sequence that has 80%
CC sequence identity to (a)-(c); or (e) a sequence that hybridises to (a)-
CC (c). Also described: (1) an expression vector comprising the above
CC nucleic acid; (2) a host cell comprising the above expression vector; (3)
CC a process for producing a polypeptide; (4) an isolated polypeptide
CC comprising: (a) an amino acid sequence encoded by any of the above
CC nucleotide sequences; (b) an amino acid sequence encoded by the full-
CC length coding region of the above nucleotide sequences; or (c) a sequence

CC having at least 80% identical to (a) or (b); (5) a chimeric polypeptide
CC comprising the above polypeptide fused to a heterologous polypeptide; (6)
CC an isolated antibody that binds to the above polypeptide; (7) a process
CC for producing the antibody; (8) an isolated oligopeptide that binds to
CC the above polypeptide; (9) a tumour-associated antigenic target (TAT)
CC binding organic molecule that binds to the above polypeptide; (10) a
CC composition of matter comprising the above (chimeric) polypeptide,
CC antibody, oligopeptide or TAT binding organic molecule, in combination
CC with a carrier; (11) an article of manufacture comprising a container and
CC the composition of matter contained within the container; (12) methods of
CC inhibiting the growth of a cell that expresses the above protein, where
CC the growth of the cell is at least in part dependent upon a growth
CC potentiating effect of the above protein; (13) a method of
CC therapeutically treating a mammal having a cancerous tumour comprising
CC cells that express the above protein; (14) a method of determining the
CC presence of a protein in a sample suspected of containing the protein
CC described above; (15) methods of diagnosing the presence of a tumour in a
CC mammal; (16) a method for treating or preventing a cell proliferative
CC disorder associated with increased expression or activity of the above
CC protein; and (17) a method of binding an antibody, oligopeptide or
CC organic molecule to a cell that expresses the protein described above.
CC The TAT sequences have cytostatic activities, and can be used in gene
CC therapy. The composition and methods are useful for diagnosing,
CC preventing or treating cancer. The composition is also used for preparing
CC a medicament for the therapeutic treatment or diagnostic detection of a
CC cell proliferative disorder or cancer. The present sequence represents a
CC human TAT cDNA sequence from the present invention.
XX
SQ Sequence 2404 BP; 590 A; 696 C; 611 G; 507 T; 0 U; 0 Other;
Query Match 34.4%; Score 658.2; DB 12; Length 2404;
Best Local Similarity 67.6%; Pred. No. 1.7e-156;
Matches 1116; Conservative 0; Mismatches 293; Indels 242; Gaps 4;
QY 2 TGCCAGGCTGATCTATCTACCTACCAACAAAGTGGCAATGCCGCTGCTGCTTC 61
DB 96 TGCCAGGCTGATCTATCTACCTACCAACAAAGTGGCAATGCCGCTGCTGCTTC 155
QY 62 TCGTCCCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 121
DB 156 TCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 215
QY 122 GTCTTCATATTCCTCTATCTATCTATCTATCTATCTATCTATCTATCTATCT 181
DB 216 ATCTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 275
QY 182 CAGCAGCTGTTTGAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 241
DB 276 CAACAGCTGATGTTGAAGAACTGATCTGATGCTGATGCTGATGCTGATGCTGCT 335
QY 242 ATATCTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 301
DB 336 ATTTCTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 395
QY 302 GAGAACTTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 361
DB 396 GAGAACTTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 455
QY 362 GATGAGAAAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 421
DB 456 GACGAGAAAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 515
QY 422 CTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 481
DB 516 CTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 575
QY 482 GTGCTACAGGTTTATGACCTATGAGGAAATGAGAGGTTGAGAGGTTGAGAGG 541
DB 576 GTGAAACAGGCTTTATGACCTATGAGGAAATGAGAGGTTGAGAGGTTGAGAGG 635
QY 542 CACAACAATTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 601
DB 636 CACAACAATTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 695

QY 602 GAGAGTGAATCCGGGACACCGTCTGCTCCAGAGCCTGAATGAGTGTACTTTG 661
 DB 696 GGAAGTGAATCCGGGACCTTATCTGCTCCAGAGCCTGTAACAGATGAACACCTTG 755
 QY 662 GGTACTGAGTGTGACAGCTGTATGCTTACAGCTGTGATTAATGCCATTTGTATACCA 721
 DB 756 CGTACTCAGTGTGACAGCTGTATGCTTACAGCTGTGATTAATGCCATTTGTATACCA 815
 QY 722 CAGTGTGTGACAGTGTGAGTATACAGTTTCTTGTGATGCTGTATGCGAGGACGTTT 781
 DB 816 CAGTGTGTGACAGTGTGAGTATACAGTTTCTTGTGATGCTGTATGCGAGGACGTTT 875
 QY 782 CTGAACCCCAACAGAGCTATCCAGAGCCTATGAGATGATCTGTGTTGCTGTATCACA 841
 DB 876 CTGAACCCCAACAGAGCCTATCCAGAGCCTATGAGATGATCTGTGTTGCTGTATCACA 935
 QY 842 ATCTGTCAATTTCTTATTTCAATGAGCTGTGAA----- 876
 DB 936 TTCTGTCAATTTCTTATTTATGTTGCTGTGAGTGTGAGCTGTCCAGAGGCTGTG 995
 QY 877 ----- 876
 DB 996 GGTGTGAGGATGCGCAGAGGGGTATGAGCAGAGCTGTGAGACGAGATGAGTGT 1055
 QY 877 ----- 876
 DB 1056 CAGGAAGAGAGGTCTCAGGGTAGAAGCAGCAGGCGTGTGCGGACACTGTATATC 1115
 QY 877 -----GG 878
 DB 1116 CCAAGTCTCGGAGAGCTGAGGACAGAGATGCTTGAACCGGAGGCGAGTGTG 1175
 QY 879 TGGCAGAACAGCTCATCAACCCCTTCCGGGAGAGCATGATTTTGAAGTCTAAGTGA 938
 DB 1176 TGGCAGAGAGCTCATCAACCCCTTGGAGAGATGATGATTTTGAAGTCTAAGTGA 1235
 QY 939 TCAATTGACGAACCTGCAAGTGTCTCCGTGTGTCCGTGTGTGAGATGACCAAGACTTGC 998
 DB 1236 TTGTGACAGGAATTTGCAAGTGTCTCCGTGTGTGAGATGAGATGACCAAGACTTGC 1295
 QY 999 CTCCCATGGAACGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1058
 DB 1296 CTCCCATGGAACGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1355
 QY 1059 CTTCGCGAGGTCTTCCGCGGATTTCTTCAATGAGCTGACCTTCAACATGACCTTAA 1118
 DB 1356 CTTCGCGCGAGGTCTTCCGCGGATTTCTTCAATGAGCTGACCTTCAACATGACCTTAA 1415
 QY 1119 AAGAAAGTGTGAGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1178
 DB 1416 AAGAAAGTGTGAGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1463
 QY 1179 ATAGACAGACCATAGAGTGTCTTCTTGAAGCTGCAACCAAAATACCATCTTCCCTTGA 1238
 DB 1464 ACGGTGATATATTGGCGCTTCTTGAAGCTGCAACCAATGATCATCTTCCAGAGG 1523
 QY 1239 AAGCTTAAAGCAAACTATTGTGTCTTAAAG-----CCCTCTCTTGAAGGCGCAT 1292
 DB 1524 CAAACTCAAGAGCAAACTACTGTGGCCCAAGAGGAAATCCTTCTTCAAGAGGCGCTGC 1583
 QY 1293 GTAAGATGCCAACCAAGAAAACAGAAAGAT-----GTCTTGA 1331
 DB 1584 CCAAAACCAACAGAGCAGCCCAAGAGCTTGAAGGCGCAGGAAGCAACAGAGCTTGA 1643
 QY 1332 AATTAAAGGTCTGAGCTTCTTGAATGTGTCTTCAAGTCTTAAAGAGAGAGCTCCCAT 1391
 DB 1644 AGCTTAAAGGTCTGAGAGCTTCAAGTCTTCAAGTCTTCAAGAGAGAGCTTCAAG 1703
 QY 1392 GTGAGCCCAAGAGCAGCCAGAGACCACTTAC 1422
 DB 1704 GTGAGCCCAAGAGCAGCCAGAGCAGCCACTCC 1734

RESULT 5
 AD083262
 ID AD083262 standard; cDNA; 2404 BP.
 XX
 AC AD083262;
 XX
 DT 07-OCT-2004 (first entry)
 XX
 DE Human tumour-associated antigenic target (TAT) cDNA sequence #76.
 XX
 KW human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;
 KW cancer; cell proliferative disorder; gene; ss.
 OS Homo sapiens.
 XX
 PN W02004060270-A2.
 XX
 PD 22-JUL-2004.
 XX
 PF 15-OCT-2003; 2003WO-US029126.
 XX
 PR 18-OCT-2002; 2002US-0418988P.
 XX
 PA (GETH) GENENTECH INC.
 PA (WU) WU T D.
 PA (ZHOU) ZHOU Y.
 XX
 FI Wu TD, Zhou Y;
 XX
 DR WPI; 2004-534300/51.
 XX
 PT New nucleic acid molecule and encoded polypeptide, for diagnosing,
 PT preventing or treating cell proliferative disorders such as cancer.
 XX
 PS Claim 1; SEQ ID NO 76; 5504dp; English.
 XX
 The present invention describes an isolated tumour-associated antigenic
 CC target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide
 CC sequences (see SEQ ID NO.1 to 4622); (b) the full-length coding region of
 CC (a); (c) the complement of (a) or (b); (d) a sequence that has 80%
 CC sequence identity to (a)-(c); or (e) a sequence that hybridizes to (a)-
 CC (c). Also described: (1) an expression vector comprising the above
 CC nucleic acid; (2) a host cell comprising the above expression vector;
 CC a process for producing a polypeptide; (4) an isolated polypeptide
 CC comprising: (a) an amino acid sequence encoded by any of the above
 CC nucleotide sequences; (b) an amino acid sequence encoded by the full-
 CC length coding region of the above nucleotide sequences; or (c) a sequence
 CC having at least 80% identical to (a) or (b); (5) a chimeric polypeptide
 CC comprising the above polypeptide fused to a heterologous polypeptide; (6)
 CC an isolated antibody that binds to the above polypeptide; (7) a process
 CC for producing the antibody; (8) an isolated oligopeptide that binds to
 CC the above polypeptide; (9) a tumour-associated antigenic target (TAT)
 CC binding organic molecule that binds to the above polypeptide; (10) a
 CC composition of matter comprising the above (chimeric) polypeptide,
 CC antibody, oligopeptide or TAT binding organic molecule, in combination
 CC with a carrier; (11) an article of manufacture comprising a container and
 CC the composition of matter contained within the container; (12) methods of
 CC inhibiting the growth of a cell that expresses the above protein, where
 CC the growth of the cell is at least in part dependent upon a growth
 CC potentiating effect of the above protein; (13) a method of
 CC therapeutically treating a mammal having a cancerous tumour comprising
 CC cells that express the above protein; (14) a method of determining the
 CC presence of a protein in a sample suspected of containing the protein
 CC described above; (15) methods of diagnosing the presence of a tumour in a
 CC mammal; (16) a method for treating or preventing a cell proliferative
 CC disorder associated with increased expression or activity of the above
 CC protein; and (17) a method of binding an antibody, oligopeptide or
 CC organic molecule to a cell that expresses the protein described above.
 CC The TAT sequences have cytostatic activities, and can be used in gene
 CC therapy. The composition and methods are useful for diagnosing,
 CC preventing or treating cancer. The composition is also used for preparing
 CC a medicament for the therapeutic treatment or diagnostic detection of a

CC cell proliferative disorder or cancer. The present sequence represents a
CC human TAT cDNA sequence from the present invention.
XX

Sequence 2404 BP; 590 A; 696 C; 611 G; 507 T; 0 U; 0 Other;

Query Match 34.4%; Score 658.2; DB 13; Length 2404;
Best Local Similarity 67.6%; Pred. No. 1.7e-156;
Matches 116; Conservative 0; Mismatches 293; Indels 242; Gaps 4;

```
QY 2 TGGCAAGCATGATCATACCTACCAACAAAGTACCAATGCCCCGCTGGTTCGTC 61
Db 96 TGGCGGCGATGACATCATCTTACCAAGCCAGGCTTAATGCCGCTTAGGCTCCCTC 155
QY 62 TCGTCCCTCTCTGTGTGGGAGAGCATCTTCAAGCTGCTGTATAGAAATTCCT 121
Db 156 TCCCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 215
QY 122 GTCTTCAATATTCCTTACTATTCATCCGTGACTCTACAGAAATGTTCTCTGAGTAT 181
Db 216 ATCTTCTGTCTGTCTACTATCATCCGCTTATTTATAGGCTGCGCTTACCGAAGAA 275
QY 182 CAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 241
Db 276 CAAGAGCTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 335
QY 242 ATATCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 301
Db 336 ATTTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 395
QY 302 GAGAACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 361
Db 396 GAGAACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 455
QY 362 GATGAGAAAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 421
Db 456 GACAGAGCAAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 515
QY 422 CTCATCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 481
Db 516 CTCATCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 575
QY 482 GTGTAGAGAGTTTATGATCCATGGGAGATAGAGAGTTGACAGAGTTGAGGCTTACA 541
Db 576 GTGAGAGAGGCTTATGATCCGAGACAGAGAGTTGAGAGAACTAGGCTTACA 635
QY 542 CACAACACATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 601
Db 636 CACAACATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 695
QY 602 GGAGGTGGAATCCGGGACACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 661
Db 696 GGAAGTCAATTCGGGACCTTATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 755
QY 662 CGTACTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 721
Db 756 CGTACTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 815
QY 722 CAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 781
Db 816 CAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 875
QY 782 CTGAACCCAAACAGAGATACCCAGGCGATGAGATGAGTGTGTGTGTGTGTGTGTGT 841
Db 876 CTGAACCCAGCAAGGCTTACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 935
QY 842 ATCTGTCAATTTATTTATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 876
Db 936 TTCTGTGAGTTCTTCTTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 995
QY 877 ----- 876
Db 996 GGCTGAGGATGTGCGCAGAGGGGTTCATGTGCCAGAGCTGCTGTAGACGAGAGTGCATGT 1055
```

```
QY 877 ----- 876
Db 1056 CAGGAAGAGAGTGTCTACGGGTAGAAAGACAGCCAGCGGTGTGGCGCACCTGTATTC 1115
QY 877 ----- 876
Db 1116 CCAGTACTCGGAGGCTGAGGAGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1175
QY 879 TGGGAGAACAGCTATCAACCCCTTGGGAGAGAGATGATGATTTTGAAGTAACTAGGA 938
Db 1176 TGGGAGAGCACTATCAACCCCTTGGGAGAGATGATGATTTTGAAGTAACTAGGA 1235
QY 939 TCATTTGACAGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 998
Db 1236 TTGTGTGAGAGAAATTTGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1295
QY 999 CTCCGATGAGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1058
Db 1296 CTCCGATGAGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1355
QY 1059 CTTCGCGCAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1118
Db 1356 CTTCGCGCAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1415
QY 1119 AAGAACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1178
Db 1416 AAGAGAGATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1463
QY 1179 ATAGCAGACCATAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1238
Db 1464 AGCTGTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1523
QY 1239 AAGCTTAAAGACCAACTATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1292
Db 1524 CAAGCTCAAGGACCAACTATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1583
QY 1293 GTAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1331
Db 1584 CCAAAACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCA 1643
QY 1332 AATTTAAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1391
Db 1644 AGCTTAAAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1703
QY 1392 GTGGCCCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGG 1422
Db 1704 GTGGCCCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGG 1734

RESULT 6
AA221228
ID AA221228 standard; cDNA; 2429 BP.
XX
XX AA221228;
AC
XX
XX 22-NOV-1999 (first entry)
DT
XX
XX Human CGICE long form cDNA sequence.
DE
XX
XX CGICE; Best's macular dystrophy; mutation; diagnosis; detection; BMD;
KW age-related macular dystrophy; ss.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH CDS 105..1412
FT /*tag= a
FT /product= "CGICE short form protein"
XX
XX W09943695-A1.
XX
XX 02-SEP-1999.
XX
```

PF 22-FEB-1999; 99WO-US003790.
XX 25-FEB-1999; 98US-0075941P.
PR 18-DEC-1998; 98US-0112926P.
XX (MERI) MERCK & CO INC.
PA (UYUP-) UNITV UPSPALA.
XX Petukhin K, Caskey CT, Metzger M, Madelinus C;
PI WPI: 1999-540560/45.
DR P-PSDB; AA229954.
XX Human and mouse polynucleotides encoding CGICE polypeptides.
XX Claim 2; Fig 4; 67pp; English.
XX The present sequence represents the human CGICE cDNA sequence, which when
CC mutated is responsible for Best's macular dystrophy (BMD).
CC Polynucleotides encoding CGICE are useful for diagnosing whether a
CC patient carries a mutation in the CGICE gene. Normal and mutated CGICE
CC proteins are useful for identifying activators and/or inhibitors of these
CC proteins, in order to treat BMD. The CGICE gene offers a simpler and the
CC cheaper method of diagnosing BMD without the need for the presence of the
CC patient. The gene may also be useful to discovering the genetic cause of
CC age-related macular dystrophy
XX
XX Sequence 2429 BP; 614 A; 694 C; 613 G; 508 T; 0 U; 0 Other;
SQ
Query Match 34.4%; Score 658.2; DB 2; Length 2429;
Best Local Similarity 67.6%; Pred. No. 1.7e-156;
Matches 1116; Conservative 0; Mismatches 293; Indels 242; Gaps 4;
QY 2 TGCAGAGCATGACTATCATCCTACACAAACAAAGTAGCCATGCCCGCTCGTTGCTC 61
DB 96 TGCCTGGCCATGACCATCATCTTACACAGCAAGTGGCTAATGCCCTTAGGCTCTTC 155
QY 62 TCGTCCCTCTCTGCTGCTGCGAGGAGCATCTAACAGCTGTGTATGAGAAATTCCTT 121
DB 156 TCCCGCCTGCTGTGTGCTGGCGGAGCATCTAACAGCTGTATATGAGAGTTCCTTA 215
QY 122 GTCTTCATATTCCTCTATCTATCTTCATCCGTGAGACTCTACAGAAATGTTCTCGAGTAT 181
DB 216 ATCTTCCTGCTGTCTATCTATCATCTCCGCTTATTTATTAAGCTGGCCCTTACGAGAA 275
QY 182 CAGAGCTGTGTGTGAGAGCTGCTCTGTACTGCGACAGCTACATTCAGCTCATCCCT 241
DB 276 CAACAGCTGATGTTTGAAGAACTGACTGTATATGCGACAGCTACATCCAGCTCATCCC 335
QY 242 ATATCTCTCTCTGCTGCTGCTTCTATGTTACATTTGTGTGAGCCGCTGTGAGCCAGTAC 301
DB 336 ATTTCTCTCTGCTGCTGCTTCTATGTTACGCTGTGTGAGCCGCTGTGAGCCAGTAC 395
QY 302 GAGAACTTGCCTGTCGCGCCGACCGGCTCATGATCCAGGCTGTAGCTTCGTGAGAGGAG 361
DB 396 GAGAACTTGCCTGTCGCGCCGACCGGCTCATGAGCTGTGTGAGGCTTCGTGAGAGGAG 455
QY 362 GATGAGAGAGCCGCTTGTGCTGCGCGACGCTCTATCCGCTACGCGCATCTCGGCGCAAGT 421
DB 456 GACGAGAGAGCCGCTGCTGCTGCGCGACGCTCATCCGCTACGCGCACTCGGCGCAAGT 515
QY 422 CTCTATCTGCGAGCATCTACGACTCGGTCTACAGAGGCTTCCCACTCTTACACACTG 481
DB 516 CTCTATCTGCGAGCATCTACGACTCGGTCTACAGAGGCTTCCCACTCTTACACACTG 575
QY 482 GTGTAGACAGTTTATGACCCCATGAGGAGACATTAAGAGTTTGCAGAGTTGGGCTTACCA 541
DB 576 GTGTAGACAGCTTTATGACTCCGCGAGAACACAGAGTTTGAAGAACTGAGCTTACCA 635
QY 542 CACAACATCTTGGGTGCTCGGTGTGTGTGTTGCCAATTGTCAATGAAGGCTTATCTT 601
DB 636 CACAACATCTTGGGTGCTCGGTGTGTGTGTTGCCAATTGTCAATGAAGGCTTATCTT 695

QY 602 GGAGTGCAGATCCGGGACACCGCTCTGCTCCAGAGCTGTATGAATGAGTGTACTTTG 661
DB 696 GGAGTGCAGATCCGGGACACCGCTCTGCTCCAGAGCTGTATGAATGAGTGTACTTTG 755
QY 662 CGTACTGATGTGAGCAGCTGTATGCTTACGACTGTATGAATGAGTGTACTTTG 721
DB 756 CGTACTGATGTGAGCAGCTGTATGCTTACGACTGTATGAATGAGTGTACTTTG 815
QY 722 CAGGTGTGACAGTGTGAGATATCAGCTTTTCTCTGATGCTTGAATCGGAGGCACTTT 781
DB 816 CAGGTGTGACAGTGTGAGATATCAGCTTTTCTCTGATGCTTGAATCGGAGGCACTTT 875
QY 782 CTGAACCCAAACAGAGCTACCCAGGCGATGAGTGTGCTGTGCTCTTACCA 841
DB 876 CTGAACCCAGCAAGGCTTACCCAGGCGATGAGTGTGCTGTGCTCTTACCA 935
QY 842 ATCTGCAATTTCTTATTTCTACATGAGGCTGCTGA 876
DB 936 TTCTGAGATTTCTTCTTATGTTGCTGTGAAGTGGGCTTTCAGAGGCTGCTG 995
QY 877 876
DB 996 GCGTGAAGCATGGCCAGAGGGCTCATGGCCAGACGCTCTTGAAGCAGAGATGACTGT 1055
QY 877 876
DB 1056 CAGGAAGAAAGGTCTCACGGGTAAAGAGCAGCGCGTGTGGCCACACCTGTATTC 1115
QY 877 878
DB 1116 CAGCTACTGGAGAGCTGAGGAGAGATTCCTTGAACCCGAGAGCGAGAGTGTGG 1175
QY 879 TGGCAGAACAGCTCATCAACCCCTTGGGAGAGAGCATGATGATTTTGAAGCTTACG 938
DB 1176 TGGCAGAGAGCTCATCAACCCCTTGGAGAGATGATGATTTTGAAGCTTACG 1235
QY 939 TCAATTGACAGAACTGACAGTGTCCCTGTGCTGTGAGTGTGACACAGAACTTGC 998
DB 1236 TTGTGACAGAGAAATTGACAGTGTCCCTGTGCTGTGAGTGTGACACAGAACTTGC 1295
QY 999 CTCCCATGAGAGCTGACATGATGATGAGAGAGAGAGCTTACGCGGCTTACAGCTG 1058
DB 1296 CTGAGATGAGAGCGGACATGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAG 1355
QY 1059 CTTCGACAGGTCTCCCGGCACTTCTTCAATGAGGCTTCACTTCAATGAGGCTTAAAG 1118
DB 1356 CTTCGACAGGTCTCCCGGCACTTCTTCAATGAGGCTTCACTTCAATGAGGCTTAAAG 1415
QY 1119 AAGAGACCTTGAAGCTTTGCTCAAAAGAGAGGCTGACACGATTAAGAGAGAGTGCCT 1178
DB 1416 AAGAGACCTTGAAGCTTTGCTCAAAAGAGAGGCTGACACGATTAAGAGAGAGTGCCT 1463
QY 1179 ATAGAGACCAATAGGCTGCTTCTTGAAGCTGCAACCCAAACCTTCAATGAGGCTTAA 1238
DB 1464 AGCTTGAATCATTTGAGGCTGCTTCTTGAAGCTGCAACCCAAACCTTCAATGAGGCTTAA 1523
QY 1239 AAGCTTAAAGACCAATATTTGTTCTTAAAG 1292
DB 1524 CAATCTCAAGAGACCAATATTTGTTCTTAAAG 1583
QY 1293 GTTAAAGTGCACCAAGAAACCAAGAAAT 1331
DB 1584 CCAAAACCAAGGAGGAGCAAGAAAGTAAAGGAGGAGAGAGCAAGAGGCTTGA 1643
QY 1332 AATTAAAGGCTGAGCTTCTTGAAGTGTCTCAAGTAAAGGAGAGAGGCTTCAAT 1391
DB 1644 AGCTTAAAGGCTGAGCTTCTTGAAGTGTCTCAAGTAAAGGAGAGAGGCTTCAAT 1703
QY 1392 GTGGCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1422
DB 1704 GTGGCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1734

RESULT 7
AB061177
ID AB061177 standard; cDNA; 2112 BP.
XX
AC AB061177;
XX
DT 26-FEB-2003 (first entry)
XX
DE FLJ20132 file clone COL6441 encoding sequence.
XX
XX Neuroprotective; immunomodulator; cancer; chromosome 19; cytosolic;
KM anti-inflammation; gene therapy; nutritional supplement; wound; burn;
KM ulcer; Alzheimer's disease; Huntington's disease;
KM amyotrophic lateral sclerosis; autoimmune disorder; inflammation;
KM vulnerability; gene; ss.
XX
OS Homo sapiens.
XX
XX MO200231111-A2.
XX
XX 18-APR-2002.
XX
XX 11-OCT-2001; 2001WO-US027760.
XX
XX 12-OCT-2000; 2000US-00687527.
XX
XX (HXS-) HXSEQ INC.
XX
XX Tang YT, Liu C, Zhou P, Auandi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX
XX WPI; 2002-426278/45.
XX
XX N-PSDB; ABP43933.
XX
XX New polypeptides and their encoded proteins, useful as nutritional
PT sources or supplements, or in gene therapy, particularly for treating
PT wounds, Alzheimer's disease, amyotrophic lateral sclerosis, cancer or
PT inflammation.
XX
XX Claim 1; SEQ ID # 390; 357bp + Sequence Listing; English.
XX
XX The invention relates to 446 newly isolated polynucleotide sequences. The
CC activity of polynucleotides of the invention may be described as,
CC vulnerability, neuroprotective, immunomodulator, cytosolic and anti-
CC inflammatory. Compositions comprising nucleic acids of the invention are
CC useful for treating a mammalian subject, or as nutritional sources or
CC supplements. These are useful in gene therapy, particularly for treating
CC wounds, burns or ulcers, Alzheimer's disease, Huntington's disease,
CC amyotrophic lateral sclerosis, autoimmune disorders, cancer or
CC inflammation. The nucleic acids and polypeptides are also useful in
CC diagnostic and research methods. The sequences given in records AB060788-
CC AB061133 represent polynucleotides of the invention. NOTE: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 2112 BP; 382 A; 681 C; 607 G; 442 T; 0 U; 0 Other;
SQ
Query Match 27.1%; Score 519.4; DB 6; Length 2112;
Best Local Similarity 66.9%; Pred. No. 3,1e-121;
Matches 755; Conservative 0; Mismatches 371; Indels 3; Gaps 1;
QY 9 CCAATGACTATCACTACCAAAAGTACCAATGCCGCTGGCTTCGTTCTGCTCC 68
DB 190 CGATGACCGCTACCTACACAGCCGAGTGGGAAAGCCGCTTGGTGGCTTCCAGC 249
QY 69 TCTCTGCTGCTGGAGGAGGAGCATCTACAGCTGCTATGAGAAATTCCTTGCTTCA 128
DB 250 TGTCTGCTACTGTGGCTGGAGCATCTACAACTCTCTGGCGAGAGCTGCTGCTTCC 309
QY 129 TATTCCTCTACTATTCATCCGTGACTCTACAGAAATGTTCTCTGAGTATCAGCAGC 188
DB 310 TTGGATTCTTACATGAGCGGTGAGTGTGCTTACCGCTTGTGCTGACCGAAGGAGCAAGC 369

QY 189 TGTGTTGAGAAAGCTGGCTCTGTACTGCGACAGCTACATTACAGCTCATCCCTATTCCT 248
DB 370 GCTACTCTCGAAGACTGTGTATTTATTTGTGACCAAGTATGCAAGCTCATCCCTGTCTCT 429
QY 249 TCGTTCGGGGTTTCTATGTTACATTTGGTGTGAGCCCGCTGGTGAAGCCATGAGAAACT 308
DB 430 TCGTGTGGGCTTTTATGTGACGCTGGTGAACCCCTGGTGAAGCCATGACCTATGCA 489
QY 309 TGGCGTGGCCGACCGCTCATGATTCAGAGTGTACTCTGTGAGAGGCAAGATGAGG 368
DB 490 TGCCTGCGCCGAGCGCTCATGTGCTGTGGGAGCACTGTCAAGAGCGAGCAAGC 549
QY 369 AAGCCGCTTGTGCTGCGGAGCAAGCTCATGCGCTACGCTAGCTGGGCAAGTGTCTATCC 428
DB 550 GCGGCCCTCTTACCGGCGCACTCATGCGCTACGAGGCTCTCGGCGCTGCTCATCC 609
QY 429 TGCAGACATCAAGCACTCGGTCTACAGCGCTTTCCTCTTACCACTGAGTGTAG 488
DB 610 TGGCTCCGTACGACCGCGGTGTTCAAGGCTTCCCAACCATAGACCAAGTGTAGG 669
QY 489 CAGGTTTATGACCATGGGGAATATAGCATGTTCAGAAATTTGGGCTTACCAACAA 548
DB 670 CTGGGTTTATGACCCGAGAGGCGCAAGAGTTGAAACCTGAACCTCATCTACACA 729
QY 549 CATTCTGGGTTGCGCTGGGTTGTGTTGCAACTTGTCAATGAAGGCTTATGAGGTC 608
DB 730 AGTACTGGGTGCTCTGCTGTCTGTTCTCACTCGGCGGACAGGCGGAGGCGC 789
QY 609 GAATCCGGGACACCGCTCTGCTCCAGAGCTGATGAATGAGTGTGATCTTGGCTATC 668
DB 790 GCATCCGGAACAAGGCGCTTAAAGCTGCTCGAGAGCTAAATGTTTTTGGGGCA 849
QY 669 AGTGTGACAGCTGTATGCTTACGATCGATGATATGCCATTTGCTGTACACACAGTGG 728
DB 850 AATGTGAATGCTCTTTTACATATACCTGATTAACCTGCTGTGACACGAGGTGG 909
QY 729 TGAAGGAGGAGTATACAGCTTTTCTTGGCATGCTTATGAGGAGGAGTTTGTGAAC 788
DB 910 TGACCATGCGACGTGTACAGCTTCTTGGCTTCTCATTTGTGCGCAAGTTCTGAGC 969
QY 789 CAACAAGAGTACCTACCGAGCCATGATGATGATCTGTGTGCTCTGTCCAAATCTGTC 848
DB 970 CGGCTCAGGGTTTCAAAAGACCAAGACCTGATGATGATGATGATGATGATGATGATG 1029
QY 849 AATTTCTATTTACATGGGCTGCTGAAGTGGCAGAAAGCTATCAACCCCTTGGGG 908
DB 1030 AGTCTTCTTCTACGCGGCTGGCTCAAGGATGCTGAGCTCATCAACCCCTTGGAG 1089
QY 909 AGGAGCATGATGATTTTGAAGCTACATGATCATTTGACAGAAAGCTGAGGTCCTCT 968
DB 1090 AGGAGCATGATGATTTTGAAGCTACATGATCATTTTGAAGTATGAAATTTCCAGGTGTCATG 1149
QY 969 TGTCTGATGATGAGATGACCAAGAACTTGTCTCCATGAGAACGTGATGATGAGAACG 1028
DB 1150 TGGCAGTGAAGAGATATGATGATGATGATGATGATGATGATGATGATGATGATG 1209
QY 1029 AGGAGCGCTCAGCGGCTTACAGAGCTGCTTGTG---CCAGGTCTGCGGCGATTCCT 1085
DB 1210 CAGCGAGGCTCGGCGCCCATACACAGCGCTATGCTTCTCACTGCGGAGCGCTTCT 1269
QY 1086 TCATGGGCTCCACTTCAACATCAGCCTTAAGAAAGAAAGAAAGCTTGAAGCT 1134
DB 1270 TCAGGAGCTCCACTTTGATCATGCTGCGCCAAAGAAAGATCAAGTT 1318

RESULT 8
ABZ80973
ID ABZ80973 standard; DNA; 2137 BP.
XX
XX ABZ80973;
XX
DT 15-OCT-2003 (first entry)

DE	XX	Human dystrophin (vitelliform macular dystrophin 2-like protein 1) gene.
XX	XX	
XX	XX	Bestrophin, vitelliform macular dystrophin 2; VMD2; ds; gene; vaccine;
KW	KW	antibesity; antidiabetic; immunomodulator; hypotensive; cardiatic;
KW	KW	antiinfective; osteopathic; antiinflammatory; cytostatic; obesity;
KW	KW	energy homeostasis; metabolism; triglyceride; body-weight regulation;
KW	KW	eating disorder; cachexia; diabetes mellitus; hypertension; gallstones;
KW	KW	coronary heart disease; hypercholesterolemia; osteoarthritis; cancer;
KX	KX	sleep apnea; chromosome 19.
OS	XX	
OS	XX	Homo sapiens.
FH	XX	
FH	XX	Key Location/Qualifiers
FT	XX	CDS 212..1741
FT	XX	/*tag= a
FT	XX	/product= "VMD2-like protein 1"
PN	XX	WO2003030922-A2.
PN	XX	
PD	XX	17-APR-2003.
PD	XX	
PF	XX	09-OCT-2002; 2002MO-EP011321.
PF	XX	
PR	XX	09-OCT-2001; 2001EP-00124059.
PR	XX	
PA	XX	(DEVE-) DEVELOPENTWICKLUNGSBIOLOGISCHE FORSCH.
PI	XX	Steveneragel A, Broenner G, Fritsch R, Eulenberg K, Ciosek T;
PI	XX	WPI, 2003-393411/37.
DR	XX	P-PSDB; ABR58056.
DR	XX	
XX	XX	New pharmaceutical compositions comprising a Bestrophin gene, polypeptide
PT	XX	or nucleic acid, for treating, alleviating and/or preventing metabolic
PT	XX	diseases, e.g. obesity, cachexia, diabetes mellitus, hypertension, or
PT	XX	gallstones.
XX	XX	
XX	XX	Claim 2; Fig 4c; 85dp; English.
XX	XX	
XX	XX	This sequence represents the coding region for a member of the human
CC	CC	bestrophin gene family designated vitelliform macular dystrophin 2 (VMD2)
CC	CC	-like protein 1. The dystrophin gene family are involved in energy
CC	CC	homeostasis and metabolism of triacylglycerides. The sequence can be used for
CC	CC	the manufacture of an agent for detecting and/or verifying, for treating,
CC	CC	alleviating and/or preventing disorders including metabolic diseases such
CC	CC	as obesity and other body-weight regulation and related disorders such
CC	CC	as eating disorder, cachexia, diabetes mellitus, hypertension, coronary
CC	CC	heart disease, hypercholesterolemia, osteoarthritis, gallstones, cancers
CC	CC	of the reproductive organs, and sleep apnea. The gene is found on human
CC	CC	chromosome 19
XX	XX	
XX	XX	Sequence 2137 BP, 399 A, 671 C, 622 G, 445 T, 0 U, 0 Other;
XX	XX	
XX	XX	Query Match 27.1%; Score 519.4; DB 8; Length 2137;
XX	XX	Best Local Similarity 66.9%; Pred. No. 3.1e-121;
XX	XX	Matches 755; Conservative 0; Mismatches 371; Indels 3; Gaps 1
QY	9	CCATGACTATCACTCAACAACAAAGTAGCCATGCCGCTGGTTCGTGTCGCC 68
DB	210	CGAATGACCGTACCACTCAACACAGCCGAGTGGCGAAGCCCGCTTCGATGAGCTTCCAGC 269
QY	69	TCCCTCTGTGCTGGGAGGAGGAGCATCTCAACAGCTGCTGTATGAGAAATTCCTGTCTTCA 128
DB	270	TGCTGCTACTGTGGGCTGGAGAGCATCTCAAACTCTCTGTGGGAGAGCTGCTTGTCTTC 329
QY	129	TATTCCTCTACTATTCATCCGTGACTCTTACAGAAATGTTCTCTCGAGTATCAGACGC 188
DB	330	TTGGGATTCATATGACGCGCTGAGTGCTGCTCAACCGCTTGTGCTACCGAAGGAGGAGAGC 389
QY	189	TGTTGTTTGAAGCTGGCTCTGTATACGCGACAGTAAATTCAGTATCCCTATATCTT 248
DB	390	GCTACTTGCAAGAGCTGTGTATTTATTTATGTACAGATATGACAGCTTCAATCCCTGTCTCT 449

QY	249	TGGTTCTGGGATTTCATATGTAATCATTTGAGTGGTACCGCTGGTGGAGCCAGTACGAAGAACT	308
Db	450	TGTGTCTTGGCTTTTATATGTAAGCGCTGTGTGTGAACCGCTGTGTGGAGCCAGTACTTATGCA	509
QY	309	TGCCGTGGCCCAACGCGCTCATGATTCAGATGTCTTACGTTCTGTGAAGGGCAAGATAGAG	368
Db	510	TGCCGTGGCCCAACGCGCTCATGTGTGTGTGGCCGGGCACTGTGCAAGGACCGGAGAC	569
QY	429	TGCGAGCATCAGACACTGATCTTCAAGCGCTTTCGCCACTCTTCCACACTGTGGTCTAG	488
Db	570	GGGGCGCGCTTCACGGCGCAACTCATGCTGTACGACAGGGCTCTTGGCGCTGCTCATCC	629
QY	630	TGCGCTCCGTACGACACCGCGGTCTTCAAGCGCTTCCCAACCATAGACCACTGTGTGAAG	689
Db	489	CAGGTTTTATGAACCATGGGGAACAATAGCAGTTGACAGATTGGGCTTACCAACAACA	548
QY	690	CTGGGTTTATGACCGCGAGGAGCCGAAGATTGAAACCTGAACTCATCTTACACA	749
Db	549	CATTCTGGGTGCGCTGGGTGTGTGTTGCAACTTGTCAAATGAAGCTTATCTTGGAGTTC	608
QY	750	AGTACTGGAGTCCGCTCGCTGTGTCTCCAACTGGCGGACACAGGCCACCGAGAGGCC	809
Db	609	GAATCCGGGACACCGCTCCGCTCCAGAGCGCTGATGAAATGAGGTGTGTATCTTGGCTATTC	668
QY	810	GCATCCGCGAACACAGCGCCCTTAAGCTGTCTGTAAGAGAGCTGTAATGTTTTCCGGGCA	869
Db	669	AGTGTGAACAAGCTGATGCTTACGACTGATGAATGATTCATTTGGTGTACACACAGGTGG	728
QY	870	AATGTGAATGCTCTTTCATATGACTGAAATTAAGGATCCCTCGTGTACAGCAGGTGG	929
Db	729	TGACAGTGGCAGTATACAGCTTTTTCCTTGCATGTCTTATGCGGAGGAGCTTCTGAAACC	788
QY	930	TGACCATGCACTGTACAGCTACTTCTGTGCTTGTCTCATTTGTGTGCGCAGTTCCTGGACC	989
Db	789	CAAAACAAGACTACCCAGGAGCATGAGATGATCTGTGTGTGCTGTCTTCAAAATCTTGC	848
QY	990	CGGCTCAGGGTTACAAACACACAGACTAGACCTGTGTGTGTGCGCAATCTTCAACCTCTTGC	1049
Db	849	AATTTTATTTCTACATGAGGGCTGTGTGAAGTGGCAGAAACAAGCTATCAACCCCTTGGGG	908
QY	1050	AGTTCTTTCTTACAGCCGCGCTGTGCTCAAGATAGTGAACAAGCTATCAACCCCTTGGAG	1109
Db	909	AGGACGATGATATTTTGAAGCTATCTGATCATTTGAAGAAACCTGACAGGTGTCCCTGT	968
QY	1110	AGGACGATGATGATCTTTGAAGCAAACTTTCATGATGAATGAAGAACTTCAAGGTGTCCATGC	1169
Db	969	TGTCCGTGATGAGGATGACACAGAACTTGTCTCCCATGGAAGGTGATCTGGAACG	1028
QY	1170	TGGCAGTGAAGAAATGATATGATGACTGTGTGTGTGTGGAAGAACTTTGATCTGGGATG	1229
Db	1029	AGGACGCGCTCAGCGCGCTTACACAAGCTGTCTTGTG---CCAGGTCTGCGCGGCAATCTT	1088
QY	1230	CAGCGGAGGCTCGCGGCCCATACACAGGGGCTATGTCTTCAGGTGCGGAGAGCTTCTCT	1289
Db	1086	TCATGGGTCCACCTTCAACATCAGCTTAAGAAAGAAAGAACTTGAAGCT	1134
QY	1290	TTCAAGGGCTCACCCTTGAACATCAGCTGTGCGCAAGAAAGACATGACGTT	1338
Db	1290	TTCAAGGGCTCACCCTTGAACATCAGCTGTGCGCAAGAAAGACATGACGTT	1338
RESULT 9			
ABZ80997			
ID	ABZ80997 standard; DNA; 1530 BP.		
XX	ABZ80997;		
XX	15-OCT-2003 (first entry)		
XX	Human bestrophin (vitelliform macular dystrophin 2-like protein 1) gene.		
XX	Bestrophin; vitelliform macular dystrophin 2; VMD2; ds; gene; vaccine;		
XX	Bestrophin; vitelliform macular dystrophin 2; VMD2; ds; gene; vaccine;		

KW antiobesity; antidiabetic; immunomodulator; hypotensive; cardiac;
KM anallipemic; osteopathic; antiinflammatory; cytosolic; obesity;
KM energy homeostasis; metabolism; triglyceride; body-weight regulation;
KM eating disorder; cachexia; diabetes mellitus; hypertension; gallstones;
KM coronary heart disease; hypercholesterolemia; osteoarthritis; cancer;
KM sleep apnea; chromosome 19.
XX Homo sapiens.
OS
XX
PH Koy Location/Qualifiers
FT CDS 1..1530
PT /tag= a
PT /product= "VMD2-like protein 1 homologue"
XX W0203030922-A2.
XX 17-APR-2003.
XX
XX 09-OCT-2002; 2002WO-EP011321.
XX
XX 09-OCT-2001; 2001EP-00124059.
XX
XX (DEVE-) DEVELOPEN ENTWICKLUNGSBIOLOGISCHE FORSCH.
XX
XX Steuernagel A, Broenner G, Fritsch R, Eulenberger K, Ciosek T;
XX WPI: 2003-393411/37.
XX P-PSDB; ABR58059.
XX
XX New pharmaceutical compositions comprising a Bestrophin gene, polypeptide
PT or nucleic acid, for treating, alleviating and/or preventing metabolic
PT diseases, e.g. obesity, cachexia, diabetes mellitus, hypertension, or
PT galactosemia.
XX
XX Disclosure: Fig 41; 85pp; English.
XX
XX This sequence represents the coding region for an alternative member of
CC the human bestrophin gene family designated vitelliform macular
CC dystrophin 2 (VMD2)-like protein 1. The dystrophin gene family are
CC involved in energy homeostasis and metabolism of triglycerides. The
CC sequence can be used for the manufacture of an agent for detecting and/or
CC verifying, for treating, alleviating and/or preventing disorders
CC including metabolic diseases such as obesity and other body-weight
CC regulation and related disorders such as eating disorder, cachexia,
CC diabetes mellitus, hypertension, coronary heart disease,
CC hypercholesterolemia, osteoarthritis, gallstones, cancers of the
CC reproductive organs, and sleep apnea. The gene is found on human
CC chromosome 19
XX
SQ Sequence 1530 BP; 257 A; 493 C; 462 G; 318 T; 0 U; 0 Other;
Query Match 27.1%; Score 519; DB 8; Length 1530;
Best Local Similarity 66.9%; Pred. No. 3.4e-121;
Matches 754; Conservative 0; Mismatches 370; Indels 3; Gaps 1;
QY 11 ATGACTATCACTACAAACAAAGTAGCCGAGCCGCTGGTGGTTCCTCCCTC 70
DB 1 ATGACCGCTACCTACACAGCCGAGTGGCAAGCCGCTTGGTGGTTCCTCCAGCTG 60
QY 71 CTCCTGCTGCGCAGGAGCAGATCTACAGAGCTGCTGATGAGAAATTCCTTCTGATA 130
DB 61 CTGCTACTGAGCGGTGGAGATCTACAACTCTGAGGAGAGCTGCTGCTTCTT 120
QY 131 TTCCTACTACTTTCATCCGTGAGCTACAGAAATGCTTCTGAGATGATCAGAGCTG 190
DB 121 GGGTTCTACATGAGGCTGAGAGCTGCTGACCTTGTGCTGACGAAAGGCGAAGCGC 180
QY 191 TTGTTTGAAGAGCTGAGCTGCTGATCTGAGAGAGCTACATTCATCCATATCCCTTC 250
DB 181 TACTTCAGAAAGCTTGTGATTTATTTGACAGATGACAGACCTCATCTCTCTTC 240
QY 251 GTTCTGGGTTTCTATGTTACATTTGATGAGAGCGCTGTTGAGACCACTACAGAACTTG 310

DB 241 GTGCTTGAGCTTTTATGACGCTGAGTGAACCGCTGAGAGCCAGTACCTATGATG 300
QY 311 CCGTGGCCCGCAGCCGCTCATGATCCAGATGCTGATTCGTGAGAGGCGAAGATGAGAA 370
DB 301 CCGCTGCCGAGCGGCTCATGTGCTGTGCTGCGGACCGTGTGACGAGCGAGAGCGC 360
QY 371 GGCCTTTGCTGCGGCGCAGCTGATCCGCTACGCGCATCTCGGAGCAAGTGTATCTG 430
DB 361 GCGCGCTCTACCGGCGCAGCATCATGCGGTACGAGAGGCTCTGCGGCTCTATCTG 420
QY 431 CGCAGCATCAGACCTGTGCTGATCAGAGCGCTTCCACTCTTCAACACCTGTGTAGCA 490
DB 421 CGCTCCGTCAGACCGGCGGCTTCAAGCGCTTCCCAACATAGACACGCTGTGAGAGCT 480
QY 491 GGTTTATGACCCATAGGAGAAATAGACAGTTGAGAGAGTGGGCTTACCAACAACA 550
DB 481 GGGTTATGACCGCGAGAGCGCAGAAAGTTGAAACCTGAACCTCATCTTACAAACAG 540
QY 551 TTCTGGGTCCTGGGCTGTGTTGCAACTTGTCAATGAAAGGCTTATCTGAGAGTGA 610
DB 541 TACTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
QY 611 ATCCGAGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 670
DB 601 ATCCGAGACCAAGCGCTTAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
QY 671 TTGAGACAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 730
DB 661 TTGAGAAATGCTTTCATGATGATGATGATGATGATGATGATGATGATGATGATG 720
QY 731 ACAATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 790
DB 721 ACATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
QY 791 AACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 850
DB 781 GCTGAGAGTGAACAAAGACCAAGCTGATGATGATGATGATGATGATGATGATGATG 840
QY 851 TTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 910
DB 841 TTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900
QY 911 GACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 970
DB 901 GACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
QY 971 TCCGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1030
DB 961 GCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
QY 1031 GCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1087
DB 1021 GCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
QY 1088 ATGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1134
DB 1081 CAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1127
RESULT 10
ADM02614
ID ADM02614 standard; cDNA; 2500 BP.
AC ADM02614;
DT 20-MAY-2004 (first entry)
XX
XX Human cDNA of the invention SEQ ID NO:1299.
DE
XX
XX 89; gene; human; gene therapy; diagnostic marker; pharmaceutical.
XX
XX Homo sapiens.
OS

EN EP1347046-A1.
XX
PD 24-SEP-2003.
XX
PF 12-APR-2002; 2002EP-00008400.
XX
PR 22-MAR-2002; 2002JP-00137785.
XX
XX (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S,
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I,
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y,
DR WPI; 2003-723558/69.
XX P-Psdb; ADM05057.
XX
PT New polynucleotides and polypeptides are useful in gene therapy, for
PT developing a diagnostic marker or medicines for regulating their
PT expression and activity, or as a target of gene therapy.
XX
PS Claim 1; SEQ ID NO 1299; 305pp; English.
XX
CC The invention relates to a novel human polynucleotide and the encoded
CC polypeptide. A polynucleotide of the invention may have a use in gene
CC therapy. An oligonucleotide of the invention ADM06202-ADM06773 is useful
CC as a primer for synthesizing the polynucleotide or as a probe for
CC detecting the polynucleotide. The polynucleotides ADM01316-ADM03758 are
CC useful in gene therapy, for developing a diagnostic marker or medicines
CC for regulating their expression and activity, or as a target of gene
CC therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides
CC are useful as pharmaceutical agents. The present sequence represents a
CC cDNA sequence of the invention.
XX
SQ Sequence 2500 BP; 689 A; 632 C; 563 G; 616 T; 0 U; 0 Other;

Query Match 23.9%; Score 458.2; DB 11; Length 2500;
Best Local Similarity 63.5%; Pred. No. 1.2e-105;
Matches 700; Conservative 0; Mismatches 403; Indels 0; Gaps 0;

QY 11 ATGACTATGACCTTACACAAAGTAGCCCAATGCCGCCCTCGGTCGTTCTGCTCCCTC 70
DB 229 ATGACTGTCACCTTACCTCAAGTAAGTAGCAATGCAACTTTTGTGATTCATAGGTTA 288
QY 71 CTCCTGTCGTCGAGGAGGAGCATCTACAGAGTGTGATGAGAAATTCCTTGTCTATA 130
DB 289 CTCCTCAAGTGAAGAGGAGCATCTACAACTACTGACAGGAAATTTATGTTTGTCT 348
QY 131 TTCCTTAATTCATTCATCCGTGACTCTACAGAAATGTTCTCTGAGATGACAGACTG 190
DB 349 GTTCTTTATACAGCAATTAAGTTTGTTGATGATTTGTTACTTACAGAGTCCAAAACGT 408
QY 191 TTGTTTGAAGAGCGGCTCTGACTGAGACGTCACATTCAGTCATCCCTATTCCTTC 250
DB 409 TACTTTGAAAATTAATATCAATTAATCTGACAGATATGCTGAAACAAATTCAGTAACCTTT 468
QY 251 GTTCTGAGTTTCATATGTTTACATTTGATGAGCGCTGATGAGACCATGAGAACTTG 310
DB 469 GTGCTTGAGTTTATGTTTACTCTGTAAGTGAACGAGTGTGAGAACAGTTTGTGAATTTG 528
QY 311 CCGTGCCCGACCGCCCTCATGATCCAGGTGTCTAGCTTCTGTGAGGGCCAAAGATGAGGAA 370
DB 529 CCGTGCCCGACAGAGGCTTAATGTTCTCATCTCTAGCAGTGTTCACGGAAGAGAGACAC 588
QY 371 GGCCGTTTGTGCGGGGCGGCGCTCATCCGTCAGCCCATCTCGGGCCAAAGTGTCTATCTG 430
DB 589 GGAGCGCTGCTTAAGAGAGCGTGAATGCGTACATCTCAATCTCCCTGCTATCTTTT 648
QY 431 CGCAGCATCAGACCTCGGCTTACAAAGCGCTTTCCCACTCTTCCACACTGATGCTAGCA 490
DB 649 CGCTCGGTGAGCACTGCTGTGTAACAAAGATTTCCCACTAGACACACGATGTTGAAGCA 708
QY 491 GGTTTATGACCCATGGGGAACATAGCAAGTTGAGAGATTGGGCTTACACACAAACACA 550

DB 709 GGTTTATGACACAGANTGAAGAAATTAATTCACACACTCAAGTCTCTCATCTGAAA 768
QY 551 TTCTGGGTGCTCCGTTGGTGTGTTGCCAATGTTGTAAGGCTTATCTTGAGGTGCA 610
DB 769 TATGGGTTCATCATCTGTTGGAAATCTTGCAATTAAGCCCGAATGAAGTGA 828
QY 611 ATCCGGGACACCGCTCTGTCAGAGCTGATGAAATGAGGTGATCTTGGGTAAGTCAAG 670
DB 829 ATCAGACAGGTGTTGATCTGCAATCAATTAATGACTGAATGATCATACCGCTTTGG 888
QY 671 TGTGACAGCTGATATGCTTACAGATGATGATATCCATTTGTTGATACACAGGTGATG 730
DB 889 TGCAGCTCTTATTCGTTATGACTGAGTTGAGATTCGCTGTTTACACCGAGTTGTC 948
QY 731 ACAATGCGATATATACAGCTTTTCTTGCATGCTTATCGGAGGAGCTTTGAAACCA 790
DB 949 ACTTGTCTGTATATACCTTCTTCTTGTGCGTGTGATGAGCGCCAGTTTGGATCCC 1008
QY 791 AACAAAGACTAACCCGACATGAGATGATCTGTGTGCTGCTGCTTCAATCTGCA 850
DB 1009 ACCAAAGCTAACCGAGGACATGACTTGTGATCTTCAATCTTCACTCCCTTACAA 1068
QY 851 TTCTTATTTCAATGCGCTGCTGAAAGTGGCAGAAACAGCTCATCAACCCCTTCGGGAG 910
DB 1069 TTCTTCTTATGAGAGATGCTTAAAGTGAAGAGAGCTTATCAACCTTTTGGAGAA 1128
QY 911 GAGCATATATATTTTGAAGATTAATCTGATATTTGAAGAAACCTGCAAGTGTCCCTGTTG 970
DB 1129 GATGATGATATATTTTGAATTAATCTGATGATTTGAAGAAATTTGAGAGGCTCTCTTTTA 1188
QY 971 TCGGTGATGGATGACACAGAACTTGCCTCCCATGGAAGTGAATGTACTGGAACGAG 1030
DB 1189 GCTGTGACAAATGACATGAGACTTACCAATATGAAGAGACATTTACTGGAGCAT 1248
QY 1031 GCAGCGCTGACCGCCCTTACACAGCTGCTTTCAGAGTCTGCGGCAATTCCTTCATG 1090
DB 1249 TGTGCTGCTGCGCCACCATFACATTTGAGAGCTGTGACTGATCATATACCTCATTTCTG 1308
QY 1091 GGCTCCACCTTCAACATCAGCCT 1113
DB 1309 GGGTCAACAGTCCAGATGGGGCT 1331

RESULT 11
AB280974
ID AB280974 standard; DNA; 2028 BP.
XX
AC AB280974;
XX
DT 15-OCT-2003 (first entry)
XX
DE Human beetrophin (vitelliform macular dystrophin 2-like protein 3) gene.
XX
KW Bestrophin; vitelliform macular dystrophin 2; VMD2; ds; gene; vaccine;
KW antiobesity; antidiabetic; immunomodulator; hypotensive; cardiant;
KW antilipemic; osteopathic; antiinflammatory; hypotensive; obesity;
KW energy homeostasis; metabolism; triglyceride; body weight regulation;
KW eating disorder; cachexia; diabetes mellitus; hypertension; gallstones;
KW coronary heart disease; hypercholesterolemia; osteoarthritis; cancer;
KW sleep apnea; chromosome 12.
XX
OS Homo sapiens.
XX
XX
FH Key Location/Qualifiers
FT 1..2028
FT CDS /tag= a
FT /product= "VMD2-like protein 3"
XX
XX MO2003030922-A2.
XX 17-APR-2003.
XX

PR 09-OCT-2002; 2002WO-EP011321.
XX
PR 09-OCT-2001; 2001EP-00124059.
XX
PA (DEVE-) DEVELOPEN ENTWICKLUNGSBIOLOGISCHE FORSCH.
XX
PI Steuernagel A, Broenner G, Fritsch R, Eulenberger K, Ciosesek T;
XX WPI; 2003-393411/37.
DR P-PSDB; ABR58057.
XX
XX New pharmaceutical compositions comprising a Bestrophin gene, polypeptide
PT or nucleic acid, for treating, alleviating and/or preventing metabolic
PT diseases, e.g. obesity, cachexia, diabetes mellitus, hypertension, or
PT galactose.
XX
XX Claim 2; Fig 4e; 85pp; English.
XX
XX This sequence represents the coding region for a member of the human
CC bestrophin gene family designated vitelliform macular dystrophin 2 (VMD2)
CC -like protein 3. The dystrophin gene family are involved in energy
CC homeostasis and metabolism of triglycerides. The sequence can be used for
CC the manufacture of an agent for detecting and/or verifying, for treating,
CC alleviating and/or preventing disorders including metabolic diseases such
CC as obesity and other body-weight regulation and related disorders such as
CC heart disease, cachexia, diabetes mellitus, hypertension, coronary
CC heart disease, hypercholesterolemia, osteoarthritis, gallstones, cancers
CC of the reproductive organs, and sleep apnea. The gene is found on human
CC chromosome 12
XX
SQ Sequence 2028 BP; 543 A; 511 C; 463 G; 511 T; 0 U; 0 Other;
Query Match 22.1%; Score 423; DB 8; Length 2028;
Best Local Similarity 62.2%; Pred. No. 9.5e-97;
Matches 697; Conservative 0; Mismatches 400; Indels 24; Gaps 1;
QY 11 ATGACTATCACTACACAAAGTAGCCCAATGCCCGCTCGTTCCTCGTCCCTC 70
Db 1 ATGACTGTCACTTCTCCAGTAAAGTACCAATCTTTTGGATTCATAGTTA 60
QY 71 CTCCTGTGCTGCGAGGAGCAGATCTTCAAGCTGTGTATGAGAAATTCCTTGTCTA 130
Db 61 CTCCTCAAGTGGAGGAGCAGATCTTCAAACTACTGTACAGGGAATTTATTTTGTCT 120
QY 131 TTCCTACTACTTTCATCCGTGAGCTCAAGAAATGTTCTCTGAGATCAGAGAGCTG 190
Db 121 GTTCTTATACAGCAATTAAGTTGGTATACAGATTGTACTTACAGGAGTCCAAATACGT 180
QY 191 TTGTTTGAAGAGCTGCTGTGTACTGACAGACTTACATTCAGCTCATCCTATATCCTTC 250
Db 181 TACTTTGAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 240
QY 251 GTTCTGGGTTTCTATGTATTAATTTGGTGTGAGCCGCTGTGTGAGCCAGTACGAACTTG 310
Db 241 GTGCTGGGTTTATGTATTAATTTCTGTGTGTGAAACCATGTGTGAAACCAATTTGTGAATTTG 300
QY 311 CCGTGGCCCGACCGCTCATGATCCAGGTGTCTAGCTTCGTGAGGAGGGAAGATGAGAA 370
Db 301 CCTGGCCGACAGGCTTAATGTCTCTCATCTCTAGCAAGTGTTCACGAGAGGAGAGCAC 360
QY 371 GAGCGTTTGTGCGGCGACGCTCATCCGCTACGCACTGTGGGCAATGTCTCATCTCTG 430
Db 361 GGGGGCCGCTTAAAGAGACGCTGATGGCTACGTAATTCACACTCCCTGCTCATCTTT 420
QY 431 CGCAGCATCAGACCTGGTCTTCAAGGCGTTTCCACTTTCACACCACTGGTGTAGCA 490
Db 421 CGCTCGGTGACACTGTGTGTACAAAGATTTCCCAATAGACACAGTGTGTGAGCA 480
QY 491 GGTTTTAAAGCAATGGGGAACATAAGAGATTGAGAGATTGGGCTTACACACACACA 550
Db 481 GGTTTTAAAGCAACAGATGAAGGAATTAATTAATTAATTAATTAATTAATTAATTAATTA 540
QY 551 TTCTGGGTGCGCTGGGTGTGTTTGCACAACTTGTCAATGAAGGCTATCTTGTGAGGTGA 610

Db 541 TATTGGTTCCATTCATCTGTGTTTGAATCTGCAATTAAGCCCGGAATGAAGTGA 600
QY 611 ATCCGGGACACCGTCCGCTCCAGAGCCGTAGTAGAGGTGTGATCTTGGTACTGAG 670
Db 601 ATCAGAGCAGGTGTGATCTGCAATCATTTAGTAGCAATGAATGATACCTGCTTGG 660
QY 671 TGTGACAGCTGTATGCTTACAGTGTAGTAAAGTATCCCATTTGTGTACACAGGTGTG 730
Db 661 TGCAGCCCTTATTCGGTTATAGACTGGGTGGGATTCGCGTGGTTTACCCAGGTGTG 720
QY 731 ACAGTGCATATACAGCTTTTCTTGTGATGCTTGTATGCGGAGCACTTTGTACACCA 790
Db 721 ACTTGTGCTATACCTTCTTGTGCGCTGATGTGACCGCAGTTTGTGATCC 780
QY 791 AACAGGACTACCCAGGCGCATAGATGATGATGCTGTGCTTCCATCAATCCGTGAA 850
Db 781 ACCAAAGCTAGGAGGCGCATGACTGTGATCTTACATTCATCCATCTTCAACCTCTAGAA 840
QY 851 TTCTTATCTCATGAGGCTGTGAG-----GTGCGAGAA 886
Db 841 TTCTTCTTCTATGAGAGATGCTTAAAGATTAATTAATTAATTAATTAATTAATTAATTA 900
QY 887 CAGCTCATCAACCCCTTTCGGGAGAGCAGATGATTTTGAAGCTTACTGATCATTTGAC 946
Db 901 CAGCTTATCAACCTTTTGGAGAAATGATGATGATTTTGAAGCTTACTGATGATGAC 960
QY 947 AGAAACCTGAGGTGTCCCTTGTGCTGTGATGAGTGAACCAAGAACTTCCCTCCCA 1006
Db 961 AGAAATTTGAGGCTCTCTTTTGTAGCTGTGACGAATGACATGATGCTTACCCAAATTTG 1020
QY 1007 GAACGTGACATGTACTGGAAGAGGAGCGGCTCAGCCGCTTACACAGCTTCTTCTGCC 1066
Db 1021 AAGAAGACATTTACTGGGAAGATTTCTGTCTCTGCGCCACATACATTTGGCAGCTCT 1080
QY 1067 AGGTCTGCGCGCATTCCTTGTATGAGCTTCAACCTTCAACAT 1107
Db 1081 GACTACTGATACCTCTATTTCTGGGGTCAACAGTCCAGAT 1121
RESULT 12
ACC59900
ID ACC59900 standard; cDNA; 1948 BP.
XX
XX ACC59900;
XX
XX 07-JUL-2003 (first entry)
XX
XX Human REMAP-13 encoding cDNA SEQ ID NO:49.
XX
XX Human; receptor and membrane-associated protein; REMAP; cytosolic;
KW antidiarrhoeal; anticonvulsant; nootropic; neuroprotective;
KW cerebroprotective; anti-HIV; anti-allergic; anti-inflammatory; cancer;
KW thymomatous; gene therapy; cell proliferative disorder; atherosclerosis;
KW neurological disorder; epilepsy; Huntington's disease; stroke; AIDS;
KW immune disorder; inflammatory disorder; allergy; developmental disorder;
KW hypothyroidism; Cushing's syndrome; infection; gene; ss.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH 333..1754
FT /*tag= a
FT /product= "REMAP-13"
PN
XX
XX W02003025130-A2.
PD 27-MAR-2003.
XX
XX 12-SEP-2002; 2002WO-US029220.
XX
XX 14-SEP-2001; 2001US-0322157P.
PR 28-SEP-2001; 2001US-0326029P.

PR 05-OCT-2001; 2001US-0327380P.
PR 12-OCT-2001; 2001US-0329198P.
PR 19-OCT-2001; 2001US-0343742P.
PR 02-NOV-2001; 2001US-0343906P.
PR 02-NOV-2001; 2001US-0343980P.
PR 16-NOV-2001; 2001US-0332426P.
PR 13-MAR-2002; 2002US-0364338P.
PR 15-MAR-2002; 2002US-0364494P.
PR 29-MAR-2002; 2002US-0369248P.
XX
PA (INCYTE GENOMICS INC.
XX
PI Warren BA, Gietzen KJ, Lai PG, Xu Y, Tran UK, Lee S;
PI Borowsky ML, Becha SD, Sanjanmala MM, Thangavelu K, Tang YT,
PI Honchel CD, Yue H, Elliott VS, Richardson TW, Azimtai Y, Chavla NK;
PI Baughn MR, Lu DM, Nguyen DB, Kalafus DP, Saperstein SK;
PI Rankum J, Lehm-Mason PM, Griffin JA, Duggan EM, Lee ST;
PI Zebardjian Y, Hatalia AJA, Gururajan R, Swarnakar A, Lee EA;
PI Marquis JP, Khare R, Emerling BM, Jhang X, Jackson AA;
XX
DR MPI; 2003-354596/33.
P-PSDB; ABR43181.
XX
PT New human receptors and membrane-associated proteins (REMAP), useful for
PT diagnosing, treating and preventing diseases or conditions associated
PT with the aberrant REMAP expression e.g. cancer, AIDS, atherosclerosis, or
PT infections.
XX
PS Claim 5; Page 224-225; 241pp; English.
XX
CC ACC59888 to ACC59923 encode the human receptor and membrane-associated
CC proteins given in ABR43169 to ABR43204 and designated REMAP-1 to REMAP-36
CC (1). (1) have cytosolic, antiarteriosclerotic, anticonvulsant,
CC neurotropic, neuroprotective, cerebroprotective, anti-HIV, antiallergic,
CC antiinflammatory and thymimetic activities, and can be used in gene
CC therapy. The REMAP polypeptides and polynucleotides are useful in
CC diagnosing, treating and preventing diseases or conditions associated
CC with the decreased expression or overexpression of REMAP, such as cell
CC proliferative (e.g. cancer, atherosclerosis), neurological (e.g.
CC epilepsy, Huntington's disease, stroke), immune/inflammatory (e.g. AIDS,
CC allergies) and developmental (e.g. hypohydroidism, Cushing's syndrome)
CC disorders, or infections. They are also useful in assessing the effects
CC of exogenous compounds on the expression of nucleic acid and amino acid
CC sequences of REMAP. The REMAPs or their fragments are useful in screening
CC compounds for effectiveness as agonist or antagonist of the polypeptides,
CC or in altering the expression of the target polynucleotide and compounds
CC that specifically bind to or modulate the activity of the polypeptide
XX
SO Sequence 1948 BP; 371 A; 622 C; 564 G; 391 T; 0 U; 0 Other;
Query Match 20.9%; Score 401.2; DB 9; Length 1948;
Best Local Similarity 61.2%; Pred. No. 3.2e-91;
Matches 711; Conservative 0; Mismatches 403; Indels 48; Gaps 2;
QY 9 CCATGACTATACCTACCAAAAGTACCAATGCCCGCTGCTTCGTCGCC 68
DB 331 CCATGACGGTTTCATACCTCTCAAGTGGCGAGGCCCGCTTGGAGTTCTCTGCC 390
QY 69 TCCCTCGTGGTGGCGAGGAGCATCTACAGCTGCTGTAGAGAAATTCCTTCTTCA 128
DB 391 TGCCTTCCTCGTGGAGGAGAAAGCATCTACAGCTCTCTACAGAAATTCCTCTTTTG 450
QY 129 TATTCCTCTACTATTCATCCGTGACTCTACAGATAGTTCCTCGAGTATCAGCAGC 188
DB 451 GGGGCTGTATGCGTGTGTAGCATCACTACCGGCTGTCTGTACCCAGAGAGAGGT 510
QY 189 TGTGTTTGAAGAGCTGCTGTATCTGCGAGAGCTACATTCACTCATTCCTTATCT 248
DB 511 ACGGTATGCTCAGAGTGGCCGCTACCTGCAACCGCTGAGAGAGCTATTCCTTGTCT 570
QY 249 TCGTTCGGGTTTCTATGTTACATTGGTGGAGCCGCTGTGAGCGACAGTACAGAACT 308
DB 571 TTGATTGGGTTTCTATGTAGCTCTCGTGTGAACCGCTGTGTCTCCAGTACAGAGCA 630

QY 309 TGCCGTGGCCGACCGGCTCATGATCCAGGTGTCTAGCTTGTGGAGGCAAGATGAGG 368
DB 631 TCCCGCTGCCAGACCACTGATGTGCTCATCTCGGCTAGGTGACAGCGCTGACACAGC 690
QY 369 AAGGCGCTTTGTGCGGCGCAGCGTCAATCCGTACAGCCATCTGGGCGAAAGTGTCAATC 428
DB 691 GGGGCGCGCTGTGCTGCGCGGACCTTCACTCCGTACCGGAACCTGTGGCTGTGGTGC 750
QY 429 TGGCAGCATCAGACCTCGGTCTACAAAGGCTTTCCCACTCTTACACCACTGGTGTAG 488
DB 751 TGCGCTGTGACGACCCGCGTCTTAAGCGCTTCCACATGAGCACTGTGTGACG 810
QY 489 CAGGTTTATGACCATGAGGGAACATAGAGATTGCGAAGTTGGGCGCTACCAACAACA 548
DB 811 CAGGTTTATGATCCCAAG 870
QY 549 CATTCGTGGTCCCTGGGTGTGTTGGCCAACTTGTCAATGAAGGCTTATCTTGAAGTTC 608
DB 871 AGTACTGGGTCCCTGGGTGTGTTGGCCAACTTGTCAATGAAGGCTTATCTTGAAGTTC 930
QY 609 GAATCCGGGACACCGCTGTGCTCAAGGCGCTGATGAATGAGGTGTATCTTGTGATCTC 668
DB 931 GAATACGTGACGATATGCTCTCTGTCTACTTTTGAAGAGCTGAACAGTACCGAGCCA 990
QY 669 AGTGTGACAGCTGTATGCTTACAGACTGATAGTATCCATTGTGTGTACACAGAGTGG 728
DB 991 AGTGCACATGCTATTTACATATGATGATGATGATGATGATGATGATGATGATGATGAT 1050
QY 729 TGACAGTGGCAGTATACAGCTTTTCTTTCGATAGCTTGTATGAGAGGAGCTTTTGAAC 788
DB 1051 TGACCAATAGCGGTCTACTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1110
QY 789 CAAGCAAGAGCTACCCAGGCGATGAGTGAATCGT----- 826
DB 1111 CAGAGCAGGAGGCTGCCAATCTGAGAGCTTGAAGCCAGGCGAGAGCCAGCCAG 1170
QY 827 -----GTGCTGTCTTCAACATCTGCAATCTTATTTCTACA 863
DB 1171 CCCTGGAGACCCGAGCATGTAGGTGCTCTCACTGCTGCACTGCTGCACTGCTGCTG 1230
QY 864 TGGGCTGTGAGAGGTGGCAGAAACAGCTCAACCCCTTGGGAGAGAGATGATGAT 923
DB 1231 CTGGCTGTGCTCAAGGTGGCTGAAACAGATCAACCATTTGTGTGAGATGAGAGACT 1290
QY 924 TTGAGATTAATCGATATGATGAGAAACCTGAGAGGTGCTGCTGCTGCTGCTGCTGCTG 983
DB 1291 TTGAGACAAATCAAGCTCATGACCGCACTTGCAAGGTGCTGCTGCTGCTGCTGCTG 1350
QY 984 TGACCCAGAACTTGCCTCCATGAGAACTGACATGTAATGGAACGAGGCGCTCAGC 1043
DB 1351 TGTACAGAAACCTTCCCGCGCTGAGAGAGCAAGTACTGAGTATGAGAGCAACCGCAGC 1410
QY 1044 CGGCTTACAC---AGTGTCTTGTGCGAGGTCTGCGGCGATTCCTTATGTGGCTCACT 1100
DB 1411 CACCTTACATGTGGGCGACCGGCGGAGGTCTGTGGGCGCTCATTCCTGCGCTCACT 1470
QY 1101 TCAACATCAAGCTTAAGAAAGA 1122
DB 1471 TCAACCTGCGCATGAGAGCA 1492
RESULT 13
AB280975
ID AB280975 standard; DNA; 1422 BP.
AC AB280975;
XX
DT 15-OCT-2003 (first entry)
XX
DE Human bestrophin (vitelliform macular dystrophin 2-like protein 2) gene.
XX
KW Bestrophin; vitelliform macular dystrophin 2; VMD2; ds; gene; vaccine;

KM antiobesity; antidiabetic; immunomodulator; hypotensive; cardiac;
KM antilipemic; osteopathic; antiinflammatory; cyostatic; obesity;
KM energy homeostasis; metabolism; triglyceride; body-weight regulation;
KM eating disorder; cachexia; diabetes mellitus; hypertension; gallstones;
KM coronary heart disease; hypercholesterolemia; osteoarthritis; cancer;
KM sleep apnea; chromosome 1.
XX
OS Homo sapiens.
XX
FH Key location/Qualifiers
FT CDS 1..1422
FT /tag= a
FT /product= "VMD-2-like protein 2"
XX MO2003030922-A2.
XX 17-APR-2003.
XX 09-OCT-2002; 2002MO-EP011321.
XX 09-OCT-2001; 2001BP-00124059.
XX (DEVE-) DEVELOPENTWICKLUNGSBIOLOGISCHE FORSCH.
XX Steuernagel A, Broenner G, Fritsch R, Eulenberg K, Ciosek T;
PI WPI: 2003-393411/37.
XX P-PSDB: ABR58058.
XX
XX New pharmaceutical compositions comprising a Bestrophin gene, polypeptide
PT or nucleic acid, for treating, alleviating and/or preventing metabolic
PT diseases, e.g. obesity, cachexia, diabetes mellitus, hypertension, or
PT galactose.
XX
XX Claim 2; Fig 4g; 85pp; English.
XX
XX This sequence represents the coding region for a member of the human
CC bestrophin gene family designated vitelliform macular dystrophin 2 (VMD2)
CC -like protein 2. The dystrophin gene family are involved in energy
CC homeostasis and metabolism of triglycerides. The sequence can be used for
CC the manufacture of an agent for detecting and/or verifying, for treating,
CC alleviating and/or preventing disorders including metabolic diseases such
CC as obesity and other body-weight regulation and related disorders such as
CC eating disorder, cachexia, diabetes mellitus, hypertension, coronary
CC heart disease, hypercholesterolemia, osteoarthritis, gallstones, cancers
CC of the reproductive organs, and sleep apnea. The gene is found on human
CC chromosome 1
XX
SQ Sequence 1422 BP; 258 A; 468 C; 413 G; 283 T; 0 U; 0 Other;
Query March 20.8%; Score 399.2; DB: 8; Length 1422;
Best Local Similarity 61.1%; Pred. No. 9.1e-91;
Matches 709; Conservative 0; Mismatches 403; Indels 48; Gaps 2;
QY 11 ATGACTATCACTACACAAACAAAGTAGCCATGCCGCTCGGTTCTGTCGTCCTC 70
Db 1 ATGACGCTTTCATACACTCTCAAGTGGCGAGGCCGCTTCGAGGTTTCTTCGCGCTG 60
QY 71 CTCCTGCTGCGGAGGAGGAGCATCTCAAGCTGTGTATGAGAAATTCCTTGTTCATA 130
Db 61 CTCCTCGCTGAGGAGGAGGAGCATCTCAAGCTCTCTCAAGAAATTCCTCTTGGG 120
QY 131 TTCCTCACTATTCATCCGTGAGCTCTACAGAAATGTTCTCTGAGTGAATCAGAGCTG 190
Db 121 GCCCTGACGCTGTGCTTAAATGATCACTACCGGCTGCTGACCCAGAGCAAGATAC 180
QY 191 TTGTTTGAAGAGCTGCTCTGACTGACAGAGTACATTACCTCATATCCCTTC 250
Db 181 GTGATGCTCAGTGCGCCGCTACTGCAACCGCTCAGAGACCTCATTCCTTGTCTTT 240
QY 251 GTTCTGAGTTTCTATGTTTACATTGGTGTGAGCCGCTGTGTGAGACCAAGAACTTG 310
Db 241 GTATTGGGTTTCTATGTAATCTCTGTGTGAAACGCTGTGTGTCCCAAGTAC 300

QY 311 CCGTGCCCGACCGCCTCATGATCCAGGTGTCTAGCTTGTGAGGCGCAAGATGAGAA 370
Db 301 CCGCTGCCAGACAGAGCTGATGTGCTATCTCGGCTAGCGTGACAGGCGTGAACAGCG 360
QY 371 GGCCTTTGCTGCGGCGCAGCTCATCCGCTACCGCATCTGGGCAAGTCTCATCTG 430
Db 361 GGCCTGCTGCGCCCGACCCCTCATCCGCTACCGCAACTGGGCTCCGCTGTGCTG 420
QY 431 CGCAGCATACGACCTCGGTCTACAGCGCTTTCCTTCACACCTGTGTGTACGA 490
Db 421 CGCTCGGTACACACCGCGCTTAAAGGCTTCCCAACATGAGACAGTGTGTGACGCA 480
QY 491 GGTTTTATGACCCATGAGGAAATAGAGAGTTGACAAAGTTGGGCTTACACAAACAA 550
Db 481 GGTTCATGTCCAGAGAAAGAGAAAGAAAGTTTGAAGCCCTGAATCCGACTTCAACAG 540
QY 551 TTCTGGGTCCTCGAGTGTGTGCTTGCACCTTGTCAATGAAGGCTTATCTTGAAGTGA 610
Db 541 TACTGGGTCCCTCGTCTGTGTGTACCACTGGCGGCCAGGCCGAGGAGGAGCGGCGA 600
QY 611 ATCCGGACACCGTCTGCTCCAGAGCTGATGATAGAGTGTGTACTTTCGTAATCAG 670
Db 601 ATACGTGACGATATCGCTCTCTGTCTACTTTTGAAGAGCTGAACAGTACCGAGCCA 660
QY 671 TGTGACAGCTGTATGCTTACAGCTGATAGTATCCATTTGTGTACACAGGTTG 730
Db 661 TGCAGCATGCTATTCATGATGATGATGATGATGATGATGATGATGATGATGATG 720
QY 731 ACAGTGCAGATATACAGCTTTTCTTCTTGCATGCTTGTATGAGGAGCACTTTCTGA 790
Db 721 ACATAGCCGCTCTACTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 780
QY 791 AACAGGATACCCAGGCGATGATGATGATGATGATGATGATGATGATGATGATGAT 826
Db 781 GAGGACGAGGCTGCAAACTCAGAAAGCTTCTGAAGCCAGGCGAGAGCCAGCCAGCC 840
QY 827 -----GTGCTGTCTTTCACAAATCTGCAATTTCTTATCTTACATG 865
Db 841 CTGGAAGACCGGACAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 900
QY 866 GCGTGTGTGAAGGTGAGGAGAACAGCTATCAACCTTCGAGGAGAGAGATGATGAT 925
Db 901 GCGTGTGTGAAGGTGAGGAGAACAGCTATCAACCTTCGAGGAGAGAGATGATGAT 960
QY 926 GAGCTATCTGATCATTTGACAGAAACCTGAGGTGCTCTTGTCTCGTGTATGAGATG 985
Db 961 GAGCAATATCAGCTCATAGACCGCACTTGCAGGTGCTCTGCTATCGTGTGAGAAATG 1020
QY 986 CACGAAACTTGCCTCCCATGAGACGTGACATGTATGAAAGAGAGGAGGCTCAGCG 1045
Db 1021 TACCAAGACCTTCCCTCCGCTGAGAGAGACAGTATGAGAGAGACCGGAGCA 1080
QY 1046 CCTTACAC---AGTGTCTTGTCCAGGTCTGCGCGGATTCCTTATAGGCTTCACCTTC 1102
Db 1081 CCTTACACTGTGTGAGAGGCGGAGGAGTCTGTGGGCTTATTCCTGAGCTTCACCTTC 1140
QY 1103 AACATCAGCTTAAAGAAAGA 1122
Db 1141 AACCTGCGCATGAGAGAGCA 1160
RESULT 14
ABL10793
ID ABL10793 standard; cDNA; 2861 BP.
XX ABL10793;
XX AC
XX DT 26-MAR-2002 (first entry)
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 26861.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 26, 2005, 07:39:17 ; Search time 1067.29 Seconds

(without alignments)
10698.156 Million cell updates/sec

Title: US-09-622-964A-28

Perfect score: 1916
Sequence: 1 gtcccaagcatgcatca.....aaaaaaaaaaaaaaaa 1916

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5552208 seqs, 297965951 residues

Total number of hits satisfying chosen parameters: 11104416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications NA:*
1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/prodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/prodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq:*
5: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
6: /cgn2_6/prodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/prodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/prodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
14: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
15: /cgn2_6/prodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
17: /cgn2_6/prodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/prodata/1/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/prodata/1/pubpna/US10_NEW_PUB.seq:*
20: /cgn2_6/prodata/1/pubpna/US11_NEW_PUB.seq:*
21: /cgn2_6/prodata/1/pubpna/US60_NEW_PUB.seq:*
22: /cgn2_6/prodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	867	45.3	1758	US-10-492-032-1	Sequence 1, Appl1
2	519.4	27.1	2137	US-10-492-032-3	Sequence 3, Appl1
3	519	27.1	1530	US-10-492-032-30	Sequence 30, Appl1
4	458.2	23.9	2500	US-10-108-260A-1299	Sequence 1299, Ap
5	423	22.1	2028	US-10-492-032-5	Sequence 5, Appl1
6	399.2	20.8	1422	US-10-492-032-7	Sequence 7, Appl1
7	227.6	11.9	1263	US-09-746-783-3	Sequence 3, Appl1
8	194.4	10.1	620	US-09-814-353-19504	Sequence 19504, A
9	187.8	9.8	1350	US-09-768-826-16	Sequence 16, Appl1
10	187.8	9.8	1350	US-10-874-484-16	Sequence 16, Appl1
11	181.2	9.5	1292	US-10-198-846-11070	Sequence 11070, A

C	12	125.2	6.5	578	13	US-10-027-632-134530	Sequence 134530,
C	13	125.2	6.5	578	17	US-10-027-632-134530	Sequence 134530,
C	14	122.2	6.4	539	16	US-10-029-386-7583	Sequence 7583, Ap
C	15	119.8	6.3	235	16	US-10-029-386-21283	Sequence 21283, A
	16	118	6.2	2914	18	US-10-723-860-7441	Sequence 7441, Ap
	17	118	6.2	2914	18	US-10-723-860-8303	Sequence 8303, Ap
	18	114.4	6.0	853	14	US-10-198-846-11346	Sequence 11346, A
	19	113.8	5.9	1954	18	US-10-723-860-7576	Sequence 7576, Ap
C	20	113.2	5.9	748	13	US-10-027-632-134529	Sequence 134529,
C	21	113.2	5.9	748	17	US-10-027-632-134529	Sequence 134529,
C	22	94.4	4.9	462	14	US-10-198-846-2561	Sequence 2561, Ap
C	23	94.4	4.9	615	14	US-10-198-846-7843	Sequence 7843, Ap
C	24	94.2	4.9	507	10	US-09-764-872-195	Sequence 195, App
C	25	75.8	4.0	3186778	13	US-10-027-632-174961	Sequence 174961,
C	26	75.8	4.0	3186778	17	US-10-027-632-174961	Sequence 174961,
	27	69	3.6	1717	17	US-10-264-237-259	Sequence 259, App
	28	66	3.4	239	18	US-10-637-855-110	Sequence 110, App
	29	66	3.4	65	10	US-09-908-975-29941	Sequence 29941, A
C	30	63.6	3.3	442	10	US-09-814-353-14343	Sequence 14343, A
	31	63	3.3	751	14	US-10-198-846-1862	Sequence 1862, Ap
C	32	62.2	3.2	260	10	US-09-814-353-1601	Sequence 1601, Ap
C	33	62.2	3.2	260	10	US-09-814-353-7959	Sequence 7959, Ap
C	34	61.6	3.2	298	10	US-09-814-353-806	Sequence 806, App
C	35	61.6	3.2	298	10	US-09-814-353-7178	Sequence 7178, Ap
C	36	60.4	3.2	463	10	US-09-814-353-13563	Sequence 13563, A
	37	56.2	2.9	3490	9	US-09-925-301-44	Sequence 44, Appl1
C	38	55.6	2.9	440	17	US-10-424-599-5477	Sequence 5477, Ap
	39	54.6	2.8	644	9	US-09-984-245-84	Sequence 84, Appl1
	40	54.6	2.8	644	10	US-09-966-262-84	Sequence 84, Appl1
	41	54.6	2.8	644	10	US-09-983-966-84	Sequence 84, Appl1
	42	54.6	2.8	644	14	US-10-059-395-84	Sequence 84, Appl1
	43	54.6	2.8	644	14	US-10-143-090-84	Sequence 84, Appl1
	44	54.6	2.8	1465	17	US-10-264-237-468	Sequence 468, App
C	45	54	2.8	1872	18	US-10-357-930-24484	Sequence 24484, A

ALIGNMENTS

RESULT 1
US-10-492-032-1
; Sequence 1, Application US/10492032
; Publication No. US20050049212A1
; GENERAL INFORMATION:
; APPLICANT: Steuermagel, Arnd
; APPLICANT: Bronner, Gunter
; APPLICANT: Fritsch, Rudiger
; APPLICANT: Eulenberg, Karsten
; APPLICANT: Ciosek, Thomas
; TITLE OF INVENTION: Bestrophin and Bestrophin Homologous Proteins Involved in the
; TITLE OF INVENTION: Regulation of Energy Homeostasis
; FILE REFERENCE: 2923-606
; CURRENT APPLICATION NUMBER: US/10/492.032
; CURRENT FILING DATE: 2004-04-08
; PRIOR APPLICATION NUMBER: PCT/EP02/11321
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: EP01124059.5
; PRIOR FILING DATE: 2001-10-09
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 1758
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-492-032-1
Query Match 45.3%; Score 867; DB 13; Length 1758;
Best Local Similarity 77.1%; Pred. No. 2.66-233;
Matches 1110; Conservative 0; Mismatches 220; Indels 39; Gaps 3;
Qy 11 ATGACTATCACTTACACAAAGTACCAATGCCGCTCGGCTTCTGTCGCTC 70
Db 1 ATGACATCACTTACACAAAGTACCAATGCCGCTCGGCTTCTGTCGCTC 60

QY	71	CTCTGTCGTCGGAGGACACATCTAAACCTGCTGATATGAGAAATTCCTGTCCTATA	130
Db	61	CTGCTGTGCTGGGGGAGACATCTAAACCTGCTAATATGGCAGTTCCTAATCTTCCTG	120
QY	131	TTTCTACTAATTCATCCGTGACCTTAACAGATGGTCTCTCAAGTATCAGACCTG	190
Db	121	CTCTGCTACTAATCAATCCGCTTTATTTAATAGCTGGCCCTCAAGAAACAAACAGCTG	180
QY	191	TTTGTTTAGAAAGCTGGGCTCTGTACCTGGCAAGCTACATTAGCTCATCCCTAATCTTC	250
Db	181	ATTTTGTAGAAATGTACTCTGTATTTGGACAGCTACATCCAGCTCATCCCATTTCCCTTC	240
QY	251	GTTCTGGGTTTCTATGTTATCATGTTGGTATGAGCCGCTGGGAGGAGCCAGTACAGAACTTG	310
Db	241	GTCTCTGGGCTTCTACGTGACGCTGCTGTGTAACCGCTGGTGGAAACATGACGAAACCTG	300
QY	311	CCGTGGCCCGACCGCCTCATATCCAGTGTCTAGCTTGTGGAGGGCAAGATGAGAA	370
Db	301	CCGTGGCCCGACCGCCTCATATGACCTGTGTGGGCTTGTGTGCAAGGACGACGAA	360
QY	371	GGCGCTTTCGTGGGGGACGCTCATCCGCTAGCCATCTGGGACCAATGCTCATCTCG	430
Db	361	GGCGGCTGCTGTGGGGGACGCTCATCCGCTAGCCAACTGGGCAACGCTCATCTCG	420
QY	431	CCGACGATCAGCACTCGGTCTAACAGCGCTTTCCACTTTTCCACCTGGTGTCTAGCA	490
Db	421	CCGACGCTCAGCACCGCAGTCTAACAGCGCTTCCGAGGCGCCAGCACTGGTGTCTAGCA	480
QY	491	GGTTTATATACCATATGGGGAACTAATGACATTTGCAGAAATTTGGGCTTACCAACAACA	550
Db	481	GGCTTTATATATCTCCGCAACAAACAGATTTGGAAATCTAGGCTTACCAACAACAATG	540
QY	551	TTCTGGGTCCTCGGGTGTGTGGTTCGCACTTGTCAATAGAGCCATCTTGGAGTGTCA	610
Db	541	TTCTGGGTCCTCGGGTGTGTGGTTCGCACTTGTCAATAGAGCCATCTTGGAGTGTCA	600
QY	611	ATCCGGGACACCGTCTGCTCCAGAGCTGTATGATAGAGTGTATCTTTCGTACTCG	670
Db	601	ATCCGGGACCCATCTGCTCCAGAGCTGTGTAAACGATGAAACACTTTCGTACTCG	660
QY	671	TGTGACAGCTGTATGCTCAAGACATGGATATATCCATTTGGTATACACAGAGTGTG	730
Db	661	TGTGACACCTGTATGCTCAAGACATGGATATATCCATTTGGTATACACAGAGTGTG	720
QY	731	ACAGTGGCAGTATTAACGCTTTTCTCTGCATGCTTATTCGAGGACAGTTTCTGAACCA	790
Db	721	ACTGTGGCGGTATTAACGCTTTTCTCTGCATGCTTATTTAGTTGGCGGACAGTTTCTGAACCA	780
QY	791	AACAGGAGCTATCCAGGGCCATGAGATGATCTGTGTGTGCTGTCTTCACAATCTCTGCA	850
Db	781	GCCAAAGGCTTATCCGCCATGATAGCTGTGACCTCTGTGTGCTGTCTTCAAGTTCTGTGAG	840
QY	851	TTCTTATTTCTACATGGGCTGTGCTGAAAGTGGCAGAAACGCTCATCAACCCCTTGGGGAG	910
Db	841	TTCTTCTTCTATGTGTGGCTGGCTGAAGGTGGAGAGAGCTCATCAACCCCTTTGGAAGAG	900
QY	911	GACGATGATGATTTTGAACATTAACCTGATCATTTGACAGAAACCTGCAGAGTGTCCCTGTG	970
Db	901	GATGATGATGATTTTGAACACCAATCGATATGTTCGACAGAAATTTCCAGAGTGTCCCTGTG	960
QY	971	TCCGTGATGAGGATGACCAAGAACTTGCTCCCATGAAACGTCATATACCTGTGAAACGAG	1030
Db	961	GCTGTGATGATGATGACCAAGAACTTGCTCCCATGAAACGTCATATACCTGTGGAATAG	1020
QY	1031	GAAGGCTTCAGCGCCCTTACACAGCTGCTTCTGCAAGTGTGCGCGGCAATTCCTTATG	1090
Db	1021	CCCAAGCCACAGCCCCCTTACACAGCTGCTTCCGCCAGTTCCTGCGAGCTCTCTTATG	1080
QY	1091	GGCTTCACCTTCAACATCAGGCTTAAAGAAATAAGACTTAAGCTTTGGTCAAAAGAGAG	1150
Db	1081	GGCTTCACCTTCAACATCAGGCTTAAAGAAATAAGAGATTCAGAGCTTCAACCTTCAAGAG	1140

QY	1151	GCTGCA CGATTAAGAAAGAGAGTGGCTATATGACGACCACTAGGCTGCTTTAGGACTG	1210
Db	1141	G-----ACGAGAGAGATGCTCACCTGGCTCATTTGGCCGCTTCTTAGGGCTG	1188
QY	1211	CAACCCAAAACTACACTCTTCCTTGAAAGACTTAAAGACCAACTATATGTGTTCTTAG	1270
Db	1189	CAGTCCCACTGATCAACATCTCTCCAGGGCAAACTCAAGAGCCAAACTACTATGTGGCCAG	1248
QY	1271	A-----ACCCCTCTCTCGAAGGCGCATGTATAGGATCCCAACGAAAAACCGAAG--	1322
Db	1249	AGGGATTCCTTCTTCACAGAGGGCTCTCCAAAAACAAGGCGCAACGAAACGATTT	1308
QY	1323	-----ATGCTGGAATTTTAGGGTCTGCACTTCTTGAATATGTT	1363
Db	1309	AGGGGCGAGAGACAACAAGGCGCTGAGAGCTTATAGCTGTGACGCGCTTCAAGTGTGC	1368
QY	1364	CCAAGGTTTAAAGAGAGAGGCTCCCACTTGTGGCCACAGGCACTCCAGACCAACCTTAC	1422
Db	1369	CCACTGTATCAGAGCCAGGCTACTATACATGAGCCCCCAAGACACCCCTCAGACCCCACTGC	1427

RESULT 2

```

US-10-492-032-3
; Sequence 3, Application US/10/492032
; Publication No. US20050049212a1
; GENERAL INFORMATION:
; APPLICANT: Steuernagel, Arnd
; APPLICANT: Bronner, Gunter
; APPLICANT: Fritsch, Rudiger
; APPLICANT: Bubenbergy, Karsten
; APPLICANT: Closser, Thomas
; TITLE OF INVENTION: Bestrophin and Bestrophin Homologous Proteins Involved in the
; TITLE OF INVENTION: Regulation of Energy Homeostasis
; FILE REFERENCE: 2923-606
; CURRENT APPLICATION NUMBER: US/10/492,032
; CURRENT FILING DATE: 2004-04-08
; PRIOR APPLICATION NUMBER: PCT/EP02/11321
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: EP01124059.5
; PRIOR FILING DATE: 2001-10-09
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 2137
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-492-032-3

```

Query Match	27.1%	Score 519.4	DB 19	Length 2137
-------------	-------	-------------	-------	-------------

Best Local Similarity 66.9%; Pred. No. 3.2e-135;

Matches 755; Conservative 0; Mismatches 371; Indels 3; Gaps 1;

QY	9	CGATGACTATACACCTACCAACAAAGTAGCCAAATGCCGGCTGGTTCGTCTCTGCC	68
Db	210	CGATGACCGCTACCTACACAGCCGAGGGGGAAGCCGCGCTTGCGTGGCTCTCCACGC	269
QY	69	TCCCTCTGCTGGCGAGCAGACATCTAAAGCTGCTGTATAGAGAAATCCCTGTCTTCA	128
Db	270	TGCTGCTACTGTGGCGTGGAGACATCTACAAATCTCTGTGGCGAGACTGCTGCTTCC	329
QY	129	TATTCCTACATATCCATCCGTGACCTTACAGATGTTCTCTCGAGTATCAGACG	188
Db	330	TTCGGTTCTACATGGGGCTGAGTGTGCTTACCCCTTTGTGTGACCGAAGGGCGAAGC	389
QY	189	TGTTGTTTGAAGCTGAGCTCTGTACTGCGACAGACTACATTCAGCTCATCCCTATATCT	248
Db	390	GCTACTTCGAGAAGCTTGATTTATTTGTGACAGTATGCGACGCTCATCTCGTGTCTC	449
QY	249	TGCTTCGAGTTCTATGTACATTGTGTGTGTGAGCCGCTGTGTGAGCCACTACAGAACT	308
Db	450	TGCTGTGTGGCTTTATGTATGTAGCGCTGTGTGTGAACCGCTGTGTGAGCCAGTACTCTATGCA	509
QY	309	TGCGGTGACCCGACCGCTCATGTATCCAGAGTGTACTCTTGTTGAGAGGCGAAGATGAGG	368

Db 510 TGCCGCTCCGACGCGCTCATGTGCGGTGGCGGGACCGTGACCGACGCGACGACC 569
Qy 369 AAGGCGCTTTGCTGGGGGACGCTCATCCGCTAGACCATCTGGGGCCAACTGCTCATCC 428
Db 570 GCGGCGGCTTACCGGGGACACATCATGCGCTAGCGGGGCTCTGGCGGTGCTCATCC 629
Qy 429 TGCCGACATGACGACCTCGGCTCTACAGCGCTTTCCACTCTTCAACCACTGGTGTAG 488
Db 630 TGCCGCTCCGCTAGACGACCGCGGTGTTCAAGCGCTTCCCAACATAGACCACTGGTGTAG 689
Qy 489 CAGGTTTATGACCCCATGGGGAACATTAAGCAGTTGACAGAGTTGGGCTTACCAACA 548
Db 690 CTGGATTATGACCGCGAGAGGCGCAGAAGTTTGAACCTGAATCATCTCTACACA 749
Qy 549 CATTTGTGGTCCCTCGGGTGTGTGTGGCACTTGTCAATGAAGCCATTTCTGGAGTC 608
Db 750 AGTACTGGGTCCCTCGCTGTGTCTTCCAACTGGCGGCAACAGCCGAGCGAGGGCC 809
Qy 609 GAATCCGGGACACCGTCTGCTCCAGAGCTGATGATGAGGTGTACTTTGGCTACTC 668
Db 810 GCATCCGGACACAGCGCCCTTAAGCTGTGCTGAGAGCTGAATGTTTTGGGGCA 869
Qy 669 AGTGTGACAGCTGTATGCTTACGACTGATTAAGTATCCATTTGGTGTACACAGGTG 728
Db 870 AATGTGAATGCTTTCACTATGACTGATTAAGCTATCCCTGTGTACACGAGGTG 929
Qy 729 TGACAGTGGAGTATACAGCTTTTCTTGTGATGTTGAGGAGGAGTTTCTGAACC 788
Db 930 TGACCATTCGACGTATACAGTACTTCTGTGGCTGTCTCATTTGTCTGCAAGTTCTGAGCC 989
Qy 789 CAACAAGAGCTACCCAGGCGCATGATGATCTGTGTGGCTGTCTTCAAACTCTGC 848
Db 990 CGGCTCAGGGTTCAAAACACACAGCTTATGACTGTGTGTGCTTCACTTCCCTTGC 1049
Qy 849 AATTCTTATTTACATGGGCTGTGCTGAAGGTGGAGAACAGCTCATCAACCTTCCGGG 908
Db 1050 AGTTCTTCTTACGCGCGGTGTGCTCAAGGTAGTGAAGCTCATCAACCTTCCGGAG 1109
Qy 909 AGAGCATGATGATTTTATGACTATCTGATCTTGAACAACCTGAGGTGTCTCT 968
Db 1110 AGGAGATGATGATCTTATGACCACTTCTGATCTGATGAACCTTCCAGGTGTCTCAATGC 1169
Qy 969 TGTCGTGATGAGGTGACACCAAGCTTCCCATGGAAGCTGACATGATCTGGAAG 1028
Db 1170 TGCGATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1229
Qy 1029 AGGAGCGCTCTAGCGCGCTTACACAGCTGTCTG---CGAGTCTGCGCGCATTTCT 1085
Db 1230 CAGCGAGGCTCTCGGCGCCCATACACAGCGGCTATCTTCCAGCTGCGGACGCTTCT 1289
Qy 1086 TCATGGGCTTCAACCTTCAACATCAGCTTAAAGAAAGAGACTTGAAGCT 1134
Db 1290 TCCAGGGCTTCCACTTTGACATCAGCTGGCCAAAGAGACATGCAAT 1338

RESULT 3
US-10-492-032-30
; Sequence 30, Application US/10492032
; Publication No. US20050049212A1
; GENERAL INFORMATION:
; APPLICANT: Steuernagel, Arnd
; APPLICANT: Bronner, Gunter
; APPLICANT: Fritsch, Rüdiger
; APPLICANT: Eulenberg, Karsten
; APPLICANT: Closssek, Thomas
; TITLE OF INVENTION: Beetrophin and Beetrophin Homologous Proteins Involved in the
; TITLE OF INVENTION: Regulation of Energy Homeostasis
; FILE REFERENCE: 2923-606
; CURRENT APPLICATION NUMBER: US/10/492,032
; PRIOR FILING DATE: 2004-04-08
; PRIOR APPLICATION NUMBER: PCT/EP02/11321
; PRIOR FILING DATE: 2002-10-09

; PRIOR APPLICATION NUMBER: EP01124059.5
; PRIOR FILING DATE: 2001-10-09
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 30
; LENGTH: 1530
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-492-032-30

Query Match 27.1%; Score 519; DB 19; Length 1530;
Best Local Similarity 66.9%; Pred. No. 3,4e-15;
Matches 754; Conservative 0; Mismatches 370; Indels 3; Gaps 1;

Qy 11 ATGACTATCACTTACACAAACAAAGTAGCCCAATGCGCCGCTGGTGTGTCGTCCTC 70
Db 1 ATGACCGTACACTACACAGCCGAGTGGCAACGCGCTTGGGTGCTTCTCCAGCTG 60
Qy 71 CTCTGTGCTGGCAGGAGCAGATCTACAGCTGTGTATGAGAAATCTTGTCTTCA 130
Db 61 CTGCTACTGTGGCGTGGAGCATCTACAACTCTGTGGGAGAGCTGTCTGCTTCTT 120
Qy 131 TTCTCTTATTCATTCATCCGTGAGACTCTACAGAAATGTTCTGTGAGTATCAGACTG 190
Db 121 GGGTTCTTACATGGGCTGAGTGTGCTGCTACCGCTTGTGTGCTGACGGAAGGCGC 180
Qy 191 TTGTTTGAAGCTGGCTGTGACTGCGAGACTTACATTCAGCTCATCCCTATTCCTTC 250
Db 181 TACTTCGAAGCTTGTGATTTATTTGACAGATGATGACAGCTCATCTGCTCTTC 240
Qy 251 GTTCTGGTTTCTATTTATTAATTTGATGAGCCGCTGTGTGAGAGCAGTAAGAACTTG 310
Db 241 GTGCTGTGCTTTATTTATGAGCTGTGTGTGAACCGCTGTGTGAGAGCATCTATGATG 300
Qy 311 CCGTGGCCGACCGCTCATGATCCAGGTGTCTTCTGTTGAGGCGCAAGATGAGGAA 370
Db 301 CCGCTGCCGACCGGCTCATGTGTGTGTGCGGCGACCGGTGACGAGCGGACGCGC 360
Qy 371 GCGCGTTTGTGGGCGGAGCTCATCGGCTACAGCACTCCGGGCAAGTGTCTACCTG 430
Db 361 GCGCGCTTACCGGCGACCTATGTGCTACGAGGAGCTCTGGCCGCTCATCTG 420
Qy 431 CGCAGCATCAGCAGCTGGTGTACAGAGCTTTCCTTCACTTCAACCTGTGTGAC 490
Db 421 CGCTCCGTCAGACCGCGGTGTTCAGAGGCTTCCCAACATAGACAGTGTGTGAGGCT 480
Qy 491 GGTTTATGACCATTTGGGAAATTAAGCATTTGAGAAATTTGGGCTTACACACACA 550
Db 481 GGGTTTATGACCCGCGAGAGCGCAAGAAATTTGAAAACCTGAATCTCATACACAAG 540
Qy 551 TTCTGGGTGCGCTGGGTGTGTGTTGCCAATTTGTCAATGAAGGCTTATCTTGGAGTGA 610
Db 541 TACTGGGTGCGCTGGGTGTGTGTTCTCAACCTTGGCGGACAGGCGGCGAGGGCGC 600
Qy 611 ATCCGGGACACGCTGCTGCTCCAGAGCTGATGATGATGATGATGATGATGATGATG 670
Db 601 ATCCGCAACACAGCGCCCTTAAGCTGTGCTCAAGAGCTGAATGTTTTCCGGGGCAA 660
Qy 671 TGTGACAGCTTATGCTTACAGTGAATGAATATCCATTTGTGTATACACAGGTGTG 730
Db 661 TGTGGAATGCTTTCATCTATGATGATGATGATGATGATGATGATGATGATGATG 720
Qy 731 ACAATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
Db 721 ACAATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
Qy 791 AACAGACTTACCAAGCAGTGAATGATGATGATGATGATGATGATGATGATGATGATG 850
Db 781 GCTGAGGTTTACAAAGACACAGACTTGAACCTGTGTGTGCTTCAACCTTCTGAG 840
Qy 851 TTCTTATTTTACATGAGGTGTGTGAGGTGTGAGAGAGTGTATCAACCTTCTGGGAG 910
Db 841 TTCTTCTTCTTACGCGGCTGCTCAAGTGTGAGAGTGTATCAACCTTCTGGAGAG 900

QY	911	GACGATGATGATTTTGA	CTAACTGATCATTTGACAGAAACCTGAGGTGCCCTTG	970
Db	901	GACGATGATGATCTTTAGACCA	CTTTCTGATGATAGAACTTCCAGGTGTCATGCTG	960
QY	971	TCGTGATGGAATGACCA	GAACTTGCTCCCATGGAACGTGACATGTA	CTGGAACGAG 103
Db	961	GCAATGGAAGAGATGAT	ATGATGACCTGGCTGCTGAGAAAGATCTTGA	CTGGAATGCA 102
QY	1031	GGAGGCGCTCAGCGCC	CTACACAGTGGTTCTG---CCAGTCTTCCGGGATTCCTTC	108
Db	1021	GGCGAGGCTCGCGCC	CCATACACAGGGCTACTGTCCTTCCAGGTGCGGACGCTTC	108
QY	1088	ATGGGCTCCACCTTCA	CACTCAAGCTTAAAGAAAGACTTGAAGCT 1134	
Db	1081	CAGGGCTCCACCTTGA	CACTCAAGCTTAAAGAAAGACTGCACTT 1127	
RESULT 4				
US-10-108-260A-1299				
; Sequence 1299, Application US/10108260A				
; Publication No. US20040005560A1				
; GENERAL INFORMATION:				
; APPLICANT: HELIX RESEARCH INSTITUTE				
; TITLE OF INVENTION: No. US20040005560A1 full length cDNA				
; FILE REFERENCE: H1-A0106				
; CURRENT APPLICATION NUMBER: US/10/108,260A				
; CURRENT FILING DATE: 2002-03-27				
; NUMBER OF SEQ ID NOS: 5458				
; SOFTWARE: Patentin Ver. 2.1				
; SEQ ID NO 1299				
; LENGTH: 2500				
; TYPE: DNA				
; ORGANISM: Homo sapiens				
US-10-108-260A-1299				
Query Match 23.9%; Score 458.2; DB 17; Length 2500;				
Best Local Similarity 63.5%; Pred. No. 6.3e-118;				
Matches 700; Conservative 0; Mismatches 403; Indels 0; Gaps 0;				
QY	11	ATGACTATCACCCTAC	CAAAACAAAGTACCAATGCCCCGCTCGTTCCTCCTC	70
Db	229	ATGACTGACCTTATCT	CAAGTAAATGACAAATGCACTTTTGTGATTCATAGCTTA	288
QY	71	CTCCTGCTGGCGAG	CAGCATCTACAAGCTGCTGTATGAGAAATTCCTTGTCTCATA	130
Db	289	CTCCTCAAGTGAAG	AGCAGCATCTACAACTACTGTACAGGAAATTTATGTTTGTCT	348
QY	131	TTCTCTTACATTCAT	CCGTGGACTTACAGAAATGTTCTCTCGAGTATCAGACGCTG	190
Db	349	GTTCTTTTATACAGA	AAATGATTTGGTGTACAGATGTTTACTTACAGAGTCCAAAACGT	408
QY	191	TTGTTTGAAGCTG	GCCTGTATCTGCGACAGCTACATTCAGCTCATCTATATCCTTC	250
Db	409	TACTTTGAAAATTA	TCAATTAATTTACTGTGACAGATATGCTGAAACAAATTTCCAGTAACCTTT	468
QY	251	GTTCTGGGTTTCTA	TATGTTACATTTGGTGTAGCGCTGTGGAGCGACATACAGAACTTG	310
Db	469	GTTGCTGGGTTTAT	ATGTTACTCTGTGTATGTACCCGATGTGTGAACCAATTTGTGTAAATTTG	528
QY	311	CCGTGGCCCGAC	CGCCCTCATATATCAAGGTGTCTAGCTTGTGAGGGCGAAGATGAGAA	370
Db	529	CCCTGGCCGACAG	CAAGCTTAATGTTCTCTCATCTTCTAGCAGTGTTCACGGAACGACGACAC	588
QY	371	GGCGGTTTGTGCG	GGCGGAGCTCATCGCTACAGCCATCCGCGGGCCAAAGTCTATCCTG	430
Db	589	GGGCGCTGCTTGA	AGAGACCTCATATGGCTACAGCTCAATCTCACCTCCCTGCTCATCTTT	648
QY	431	CGAGCATACAGCA	CTCGATCTACAAAGCGCTTTTCCACTCTTTCACCACTGAGTGTCTAGCA	490
Db	649	CGCTCGGTGACAT	GTGCTGTGTATCAAAAAGATTTTCCACAAATGACCAACGATGTGTGAAGCA	708
QY	491	GGTTTATATGACC	ATGAGGAAATATAGCAAGTTGAGAAAGTTGGGCTTACCAACAAACACA	550

Db	Query	Score	DB	Length
Db	709 GGTTTTATGACACAGATGAAAGAAATTTATTCACCACTCAAGTCTCCATCTGAAA	22.1%	DB 19	Length 2028
Query	551 TTCTGGGTGCTCCGGGTGGTGGTTTCCCACTTGCAATGAAAGCTTATCTTGAGGTGCA	62.2%	Pred. No. 4,7e-108	
Db	769 TATTTGGTTCATATTCCTGGTTTGGAAATCTTGCAACTAAAGCCCGAAATGAAGGTAGA			
Query	611 ATCCGGGACACGCTCCGCTCCAGAGCTGATGAAATGAGGTGGTGAATTTGGCGGATCTCAG			
Db	829 ATACAGACAGGTGTTGATCTGCATATCATTTATGATCTGAAATGAATGCATACCGCTTTGG			
Query	671 TGTGACAGCTGTATGCTCAAGACTGATTAAGTATCCCATTTGGTGTATACACAGGTGATG			
Db	889 TCGAGCTCTTATTCGGTATATGACTGGGATTCGGGATTCGGGTTTATACCCAGGTGTCT			
Query	731 ACAGTGCAGTATATACAGCTTTTTCCTTGCAATGCTTGATCGGGAGGCAAGTTTGAACCA			
Db	949 ACTCTTGCTGTATACCTTCTTCTTGGGTGCTGATTTGGAGCCAGTTTGGATCCC			
Query	791 AACCAAGACTTACCCAGGCCATGATGATCTGTTGTGCTGCTGCTTCAATCTGCA			
Db	1009 ACCAAGGCTAGSCAAGGCAATGACTTGATCTTACATTTCCATCTTCAACCTCTCTACAA			
Query	851 TTCTTATTTCAATGAGGCTGGCTGAGGTGGACAAAGCTCATCAACCCCTTGGGAG			
Db	1069 TTCTTCTTCTATGACAGATGAGCTTAAAGTATGACAGACAGCTTATCAACCTTTTGGAGAA			
Query	911 GACGATGATGATTTTGAAGCTAACTGATCATTTGACAGAAACCTGAGGAGTCCGTTG			
Db	1129 GATGATGATGATTTTGAATCTAATCTGTGCTATGACAGAAATTTGCAAGTCTCTTTTGA			
Query	971 TCCGTGATGAGATGACACAGAACTTGTCTCCATGAAAGTGCATGTATCTGGAACGAG			
Db	1189 GCTGTGACGAATATGACATGAGCTTACCAAGATGAAAGAGCAATTAACTGGAGCAT			
Query	1031 GGAGGGCTCAAGCCGCCATACAGAGTGTCTGTCAGAGTCTGCCGGCATTCCTTCATG			
Db	1249 TTGCTGCTCGCCACCAATACATTTGGCAGCTGTGACTACTGCAATACCTTCATTTCTG			
Query	1091 GGCTCCACCTTCACATCAGCCT			
Db	1309 GGCTCAACAGTCCAGATGGGCT			

Matches 697; Conservative 0; Mismatches 400; Indels 24; Gaps 1;

QY 11 ATGACTATACCTTACACAAAGTAGCCATGCCCGCTGGTGGTCTGCTCCCTC 70
Db 1 ATGACTATACCTTACACAAAGTAGCCATGCCCGCTGGTGGTCTGCTCCCTC 60

QY 71 CTCCTGCTGGCGAGGAGCATCTCAAGCTGCTGATGAGAAATTCCTTCTCATA 130
Db 61 CTCCTGAGTAGGAGGAGCATCTCAAGCTGCTGATGAGAAATTCCTTCTCATA 120

QY 131 TTCCTTACTATTCATCCGCTGAGCTTCAAGAAATGCTTCTGAGATGACAGCTG 190
Db 121 TTCCTTACTATTCATCCGCTGAGCTTCAAGAAATGCTTCTGAGATGACAGCTG 180

QY 191 TTGTTTGAAGAGCTGGCTGCTGAGCTGAGCAATTCAGCTCATCCCTATATCCCTC 250
Db 181 TACTTGAAGAAATTAATCAATTTACTGATACAGATATGCTGAACAAATTCAGTAACTTT 240

QY 251 GTTCTGGGTTCTATGTTACATTTGGTGTGAGCCGCTGGTGGAGCCAGTACGAGAACTTG 310
Db 241 GTGCTTGGGTTTATGTTACTCTGTGTATGAAACGATGATGAGAACAGTTGTGAAATTTG 300

QY 311 CCGTGCCCGACCGCTCATGATCCAGGTGTCTAGCTTCTGAGAGGCGAGATGAGAA 370
Db 301 CCGTGCCCGACCGCTCATGATGCTCTCATCTAGCAAGTGTTCACGGAAGCGAGAGCAC 360

QY 371 GCGCGTTTGTGGCGCGAGCTCATCCGCTAGCCATCCCTGGGCGAAGTCTCATCTG 430
Db 361 GCGCGCTGCTTAAAGAGACGCTGATGCTAGCAATTCACCTCCCTCATCTCTT 420

QY 431 CGCAGCATCAGACCTCGGTCTACAGAGCGCTTTCCCACTTCCACACTGCTGCTAGCA 490
Db 421 CGCTCGGTGAGCACTGCTGTGTACAAAGATTTCCCACTGAGCACTGCTGTTGAAGCA 480

QY 491 GGTTTATGACCCATGGGAAATAGAGATTTGAGAAATTTGAGGCTTACACACACACA 550
Db 481 GGTTTATGACCAAGATGAAAGAAATTTATCAACCACTCAAGTCTCTCATCTGAAA 540

QY 551 TTCCTGGGCTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 610
Db 541 TATTTGGGTTTCATTCATCTGTTTGGAAATCTTGCACTAAAGCCCGGAATGAAAGTGA 600

QY 611 ATCCGGGACACCGTCTGCTCCAGAGCTGATGATGATGATGATGATGATGATGATGATG 670
Db 601 ATCCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660

QY 671 TGTGAGCAGCTGATGCTGATGATGATGATGATGATGATGATGATGATGATGATG 730
Db 661 TGCAGCGCTTATTCGCTTATGATGATGATGATGATGATGATGATGATGATGATG 720

QY 731 ACAGTGGCAGTATACGCTTTTCTGATGCTGATGATGATGATGATGATGATGATGATG 790
Db 721 ACTCTTGTCTATACCTTCTTCTTGTGCGCTGATGATGATGATGATGATGATGATG 780

QY 791 AACAGAGCTACCGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATG 850
Db 781 ACCAAAGCTACCGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATG 840

QY 851 TTCTTATCTACATGGGCTGGCTGAG-----GTGGCAGAA 886
Db 841 TTCTTCTTCTATGAGATGGCTTAAAGATTAATAATATGTAATGTAATGTAATGTA 900

QY 887 CAGCTCATCAACCCCTTGGGAGAGAGATGATGATGATGATGATGATGATGATGATGATG 946
Db 901 CAGCTTATCAACCCCTTGGGAGAGATGATGATGATGATGATGATGATGATGATGATG 960

QY 947 AGAAACCGGAGGTGCTGCTGTTGTCGATGATGATGATGATGATGATGATGATGATG 1006
Db 961 AGAAATTTGAGAGGTCTCTTTTACCTGATGATGATGATGATGATGATGATGATGATG 1020

QY 1007 GAACGTGACATGATCTGAG 1066
Db 1021 AAGAGAGACATTTTACTGGAGAGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080

QY 1067 AGGTCTGCCGGCAATTCCTTATGAGCTTCAACCTTCAACAT 1107
Db 1081 GACTACTGCAATACCCCTATTTCTGGGTCAACAGTCCAGAT 1121

RESULT 6
US-10-492-032-7
; Sequence 7, Application US/10492032
; Publication No. US20050049212A1
; GENERAL INFORMATION:
; APPLICANT: Steuernagel, Arnd
; APPLICANT: Bronner, Gunter
; APPLICANT: Fritsch, Rudiger
; APPLICANT: Eulenber, Karsten
; APPLICANT: Closesek, Thomas
; TITLE OF INVENTION: Bestrophin and Bestrophin Homologous Proteins Involved in the
; FILE REFERENCE: 2923-606
; CURRENT APPLICATION NUMBER: US/10/492,032
; CURRENT FILING DATE: 2004-04-08
; PRIOR APPLICATION NUMBER: PCT/EP02/11321
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: EP01124059.5
; PRIOR FILING DATE: 2001-10-09
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7
; LENGTH: 1422
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-492-032-7

Query Match 20.8%; Score 399.2; DB 19; Length 1422;
Best Local Similarity 61.1%; Pred. No. 2e-101;
Matches 709; Conservative 0; Mismatches 403; Indels 48; Gaps 2;

QY 11 ATGACTATACCTTACACAAAGTAGCCATGCCCGCTGGTGGTCTGCTCCCTC 70
Db 1 ATGAGGTTTATATACATCTCTCAAGTGGCGAGGCGCGCTTGGAGTTTCTTGGGCTG 60

QY 71 CTCCTGCTGGCGAGGAGCATCTCAAGCTGCTGATGAGAAATTCCTTCTCATA 130
Db 61 CTCCTCGGTGAGAGGAGATCTACAGGCTCTTCAAGAAATTCCTTCTGCTGG 120

QY 131 TTCCTTACTATTCATCCGCTGAGCTTCAAGAAATGCTTCTGAGATGACAGCTG 190
Db 121 GCTTGTACGCTGTGCTTACATCCTACCGCTGCTGCTGAGCCAGAGAGATGATC 180

QY 191 TTGTTTGAAGAGCTGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 250
Db 181 GTGATGCTCAGGTGGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATG 240

QY 251 GTTCTGGGTTCTATGTTACATTTGGTGTGAGCCGCTGGTGGAGCCAGTACGAGAACTTG 310
Db 241 GTATGGGTTTCTATGATGATCTGTGTGATGATGATGATGATGATGATGATGATGATG 300

QY 311 CCGTGCCCGACCGCTCATGATCCAGGTGTCTAGCTTCTGAGAGGCGAGATGAGAA 370
Db 301 CCGTGCCCGACCGCTCATGATGATGATGATGATGATGATGATGATGATGATGATG 360

QY 371 GCGCGTTTGTGGCGCGAGCTCATCCGCTAGCCATCCCTGGGCGAAGTCTCATCTG 430
Db 361 GCGCGCTGCTTAAAGAGACGCTGATGCTAGCAATTCACCTCCCTCATCTCTT 420

QY 431 CGCAGCATCAGACCTCGGTCTACAGAGCGCTTTCCCACTTCCACACTGCTGCTAGCA 490
Db 421 CGCTCGGTGAGCACTGCTGTGTACAAAGATTTCCCACTGAGCACTGCTGTTGAAGCA 480

QY 491 GGTTTATGACCCATGGGAAATAGAGATTTGAGAAATTTGAGGCTTACACACACA 550
Db 481 GGTTCATGTCCTCAGGAAGAGGAAAGATTTGAGAGCTGGAATCCGACTTCAACAG 540

QY	551	TTCTGGGGGACCTGGGAGTGGTTGGTTCACAACTTGTCACAAAGAGGCTATCTTTGGAGGTGGA	610
Db	541	TACTGGGTCCTCCCTGGTGTCTGGTTACCAACTGGGCGCCAGAGCCCGAGGGAGCGAGCGGA	600
QY	611	ATCCGGGACACCGCTCTGCTGCTCCAGAGCCGTGAATGATGAGTGTGTACTTTGCTACTCAG	670
Db	601	ATTAGGTGACGATATGCTCTCTGCTGTACTTTTGGAAAGGCTGAACAACTACCGAGCCAAAG	660
QY	671	TGTGGACAGCTGTATGCTCTACGACTGGATTAAGTATCCATTGGTGTACACACAGTGTGT	730
Db	661	TGCAGCAGCTATTCTCACTATGACTGAGTCAGCAATCCCCCTGCTACACCCCAAGTGTGT	720
QY	731	ACAGTGGAGTATACAGCTTTTTCCTTGTCAGCTTGATCGGAGGAGGTTCGTGAACCCA	790
Db	721	ACCATAGCCGTCTACTCTTCTTTGCTTCCCTCTCCCTGGTTGGCCGCCAGTGTGTGAACCA	780
QY	791	AACAAGAGCTACCCAGAGCCATGAGATGAGATCTGGTT-----	826
Db	781	GAGGCGAGGGGCTGCCAACACTCAGAAAGTTCTGAAGCCAGGCCAGAGCCGCCCAAGCC	840
QY	827	-----GTGCTGTCTTTCACAACTCGTCAATTCTTATCTTACATG	865
Db	841	CTGGGAGACCCGAGACATGTATGCTGCTCTCACACTGCTGAGATTCTTCTCTANAGCT	900
QY	866	GGCTGTGCTGAAGGTGGCAGAAACAGTCTATCAACCCCTTCGGGAGAGACAGATGATGATTT	925
Db	901	GGCTGGCTCAAGGTGGCTGAACAGATCTCAACCCATTGGTGTGAGAGATGATGACGACTTT	960
QY	926	GAGACTAATCGATCATATGACAGAAACCTGAGAGGTGCTCCCTGTTGTCCGTGGATGGAGTG	985
Db	961	GAGACAAATCAGCTCATATGACCGCAACTTGGAGGTGTCCCTGTCTATCCGTGTGAAGAAAT	1020
QY	986	CACCAGAACTTGTGCTTCCCATGGAACGTGATCTGGAACGAGGACGCGCTTCAGCCG	1045
Db	1021	TACCAGAACCTTCCCCCGCTGAGAAAGGACCAAGTACTGGATGTAGAACCAACCCGACGCA	1080
QY	1046	CCCTTAAC--AGTGTCTTTCGCCAGGTCTGCGCGGCAATTCCTTCATNAGGCTTCACACTTC	1102
Db	1081	CCCTAACACTGTGGCCACGCGGCGGAGTCTGTGGGCCCTCATTTCTGTGGCTTCACACTTC	1140
QY	1103	AACATCAGCCTAAGAAAGA	1122
Db	1141	AACCTGCGCATGAGCGACGA	1160
RESULT 7			
US-09-746-783-3			
Sequence 3, Application US/09746783			
Publication No. US20030044935A1			
GENERAL INFORMATION:			
APPLICANT: Jacobs, Kenneth			
McCoy, John M.			
LaVallie, Edward R.			
Racie, Lisa A.			
Trecey, Maurice			
Spaulding, Vikki			
Agostino, Michael J.			
Howes, Steven H.			
Reichel, Kim			
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES			
ENCODING THEM			
NUMBER OF SEQUENCES: 231			
CORRESPONDENCE ADDRESS:			
ADDRESS: Genetics Institute, Inc.			
STREET: 87 Cambridgepark Drive			
CITY: Cambridge			
STATE: MA			
COUNTRY: U.S.A.			
ZIP: 02140			
COMPUTER READABLE FORM:			
MEDIUM TYPE: floppy disk			
COMPUTER: IBM PC compatible			
OPERATING SYSTEM: PC-DOS/MS-DOS			

```

SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/746,783
FILING DATE: 21-Dec-2000
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Milasincic, Debra J.
REGISTRATION NUMBER: 46,931
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 742-4214
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1263 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-746-783-3

Query Match      11.9%; Score 227.6; DB 10; Length 1263;
Best Local Similarity 67.7%; Pred. No. 4.8e-53;
Matches 383; Conservative 0; Mismatches 144; Indels 39; Gaps 3;

QY      884  GAAcAGCTcATcAACCCCTTGGGGGAGAGcATGATGTTTGGAGcCTAAGTGAATcATT 943
Db      1    GAGAGcGCTcATcAACCCCTTGGAGAGAGATATGATGATTTTGGAGcCAACCTGAGTTGTC 60
QY      944  GACAGAAACCTGcAGAGTGTCCCTGTGTCGTGATGGATGcACACAGAACTTGCCTCC 1003
Db      61  GACAGGAATTTGcAGAGTGTCCCTGTGCTGTGATGAGATGcACAGAGcCTGCTCGG 120
QY      1004 ATGAAACGTGcATGATGTACTGGAACAGAGcAGcGCTTACGCCCTTACAGcGTGCTTCT 1063
Db      121  ATGAGGCCGAGcATGTACTGGAATAAGCCGAGcCACAGCCCCCTTACAGcGTGCTTCC 180
QY      1064 GCCAGGTCTGCCGGcCATTTCTTATGGcCTTcACCTTCAACATcAGcCTTAAAGAAAGA 1123
Db      181  GCCAGTTCCTGcAGGCTCTCTTATGGcCTTcACCTTCAACATcAGcCTTAAAGAAAG 240
QY      1124 GACTTAGAGCTTGTGcTCAAAAGAGAGAGcGTGACcGcGATAGAAAGAGTGTGCTATGc 1183
Db      241  GAGATGAGTTCcAGCCCAATcAGGAGG-----ACGAGAGAGATGCTcAGCGT 288
QY      1184 AGcACcATAGcGCTcCTTcTAGAcTcGACcTcGAAcCCAAAACTAACAATCTTCCCTTGAAGAc 1243
Db      289  GGCATcATTTGGcCGCTTCTcTAGGcCTGcAGTcCCATGATcACATcCTCCcAGGcGCAAAc 348
QY      1244 TTAAAGACCAAACTATTGTGTTCTTAAGAA-----CCCCCTCTCTGAAAGcCAcGTGTAG 1297
Db      349  TCAAGAGcACAAACTACTGTGcGCCcCAAGAGGGAATcCTTCTTCCAGAGGcGCTGCCCAA 408
QY      1298 GATGCCAACcAGAAAAACcAGAAAGAT-----GTCTGGAATATT 1338
Db      409  AACCAcAAGcGAGcCAAAcAGAAAGTTAGGGcCAcGAGAGAcCAACAGcGCTTGAAGcTT 468
QY      1337 AAGGcTcTGAcCTTCTTGAATGTGTTCCAAAGTTTAAAGAGAGcGCTCCcATTGTGTGc 1396
Db      469  AAGGcTGTGAGcGcCTTCAAGTCTGcCCcCATGTATcAGAGcCAcAGcCTAATcAAGTGTGC 528
QY      1397 CCACAGGcCAcCCAGcAGcCAcCCCTAGc 1422
Db      529  CCACAGAGcGcCCCTcAGcCCcCAcCTGC 554

RESULT 8
US-09-814-353-19504/c
; Sequence 19504, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela

```

APPLICANT: Lillie, James
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
FILE REFERENCE: MRI-006B
CURRENT APPLICATION NUMBER: US/09/814,353
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: US 60/191,031
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: US 60/207,124
PRIOR FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: US 60/211,940
PRIOR FILING DATE: 2000-06-15
PRIOR APPLICATION NUMBER: US 60/216,820
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: US 60/220,661
PRIOR FILING DATE: 2000-07-25
PRIOR APPLICATION NUMBER: US 60/257,672
PRIOR FILING DATE: 2000-12-21
NUMBER OF SEQ ID NOS: 22037
SOFTWARE: PastSeq for Windows Version 4.0
SEQ ID NO 19504
LENGTH: 620
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1
OTHER INFORMATION: n = A,T,C or G
US-09-814-353-19504

Query Match 10.1% Score 194.4; DB 10; Length 620;
Best Local Similarity 63.9%; Pred. No. 7.5e-44;
Matches 294; Conservative 0; Mismatches 166; Indels 0; Gaps 0;

QY 9 CCATGATATACCTTACCAACAAGTGGCCGCTGCTGTTCTGTCCTCC 68
DB 464 CCATGACGGTTTCAATACCTCAAGTGGCGAGGCCGCTTGGAGTTCTCTGCC 405
QY 69 TCCCTCTGTGTGGCGAGGACATCTACAGCTGCTGTATGAGAAATTCCTTGTCTCA 128
DB 404 TGCTTCTCCGCTGAGGGGAAGCATCTACAGCTCTCTACAGGAATTCCTCTCTTTG 345
QY 129 TATTCCTCTATTTCCATCCGTGAGCTCTACAGATGTTCTCGAGTATGAGCAGC 188
DB 344 GGGCTTGTACCTGTGTGATGACATCACTACCGCTGCTGCTGACCAAGAGCAGGT 285
QY 189 TGTGTTGAGAGCTGCTGTGATGCGACAGCTACATTCAGCTCATCCCTATATCT 248
DB 284 ACGTATGCTCAGGTGCGCGTACGCAACCGCTACGAGACCTCATTCCTTGTCT 225
QY 249 TCGTTCTGGGTTTCTATGTTACATTTGTTGTTGAGCCGCTGTGAGCCAGTACGAACT 308
DB 224 TTGTATGGGTTTCTATGTTGATCTCTCGTGTGAACCGCTGTGTGCTCCAGTACAAAGCA 165
QY 309 TGGCGTGGCCGACCGCTCATGATCCAGTGTCTAGTTGTGTGAGGCAAGATGAGG 368
DB 164 TCCCGTGCACAGCAGTATGTCGCTCATCTCGAGCTAGCGTACCGCTGAGCAGC 105
QY 369 AAGCCGTTTGTCTGCGGCGACAGCTCATCCGCTACGCGCATCTGGGCGAAATGCTCATCC 428
DB 104 GGGGCGGCTGCTGCGCGGACCGCTCATCCGCTACGCGAACTGGCGTTCTGTGTGTGC 45
QY 429 TGGCGAGCATAGCACTCGGTCTACAAAGCGCTTTCCAC 468
DB 44 TGGCTCGGTACGACCGCGTGTAGCGCTTCCAC 5

RESULT 9
US-09-768-826-16
Sequence 16, Application US/09768826
Patent No. US20020012966A1
GENERAL INFORMATION:

APPLICANT: Shi et al.
TITLE OF INVENTION: 18 human secreted proteins
FILE REFERENCE: PF512P1
CURRENT APPLICATION NUMBER: US/09/768,826
CURRENT FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: PCT/US00/22350
PRIOR FILING DATE: 2000-08-15
PRIOR APPLICATION NUMBER: 60/148,759
PRIOR FILING DATE: 1999-08-16
NUMBER OF SEQ ID NOS: 61
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 16
LENGTH: 1350
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (1135)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1148)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1166)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1174)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1181)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1229)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1209)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1256)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1287)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1290)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1285)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1305)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1324)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1339)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1341)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1343)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1345)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1348)
OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: SITE
LOCATION: (1348)
OTHER INFORMATION: n equals a,c,t,g, or c
US-10-874-484-16

Query Match 9.8%; Score 187.8; DB 18; Length 1350;
Best Local Similarity 60.1%; Pred. No. 8.3e-42;
Matches 377; Conservative 0; Mismatches 202; Indels 48; Gaps 2;

```
QY 544 CAACACATTCGGGGTCCCTGGGTGGTGGTCCCACTGTCATGAAAGGCTTACTTGG 603
DB 58 CAACAGACTGGGGTCCCTGGGTGGTGGTCCCACTGTCATGAAAGGCTTACTTGG 117
QY 604 AGCTGCAATCCGGGACACCGCTGCTGCCTCAGAGCCTGATGATGATGATGATGATG 663
DB 118 CGGGCGAATAGTGAACCATTCGCTCTGTCTACTTTGGAAAGCTGAAACAGTACCG 177
QY 664 TACTCAGTGGACACTGTATGCTTACACTGTGATTAAGTATCCCATTTGTGTACACA 723
DB 178 AGCCAGTGCAGATGCTATTCACATGATGATGATGATGATGATGATGATGATGATG 237
QY 724 GGTGTGACAGTGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 783
DB 238 AGTGTGACATGACCGCTGCTACTCTTCTTGGCCCTGCTGCTGCTGCTGCTGCTGCT 297
QY 784 GAACCAAAACAAGGACTACCCAGGCGATGATGATGATGATGATGATGATGATGATGATG 826
DB 298 GAGAGCCAGAGGAGGAGGCTGCCAACTCAGAACTTTGTAAGCCAGGCGAGAGCCAGC 357
QY 827 -----GTGCTGTCTTCAAACTCTGCAATCTTCAATCTTCAATCTTCAATCTTCAAT 858
DB 358 CCCAGCCCTGGGAGACCCGAGCATGTATGATGATGATGATGATGATGATGATGATGATG 417
QY 859 CTAATGAGGCTGGCTGAAAGGTGGCAAGACACTCATCAACCTTCCGCGAGAGAGATGA 918
DB 418 CTAATGAGGCTGGCTGAAAGGTGGCAAGACACTCATCAACCTTCCGCGAGAGATGA 477
QY 919 TGATTTGAGACTAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 978
DB 478 CGACTTTGAGCAAAATGCTCATGATGATGATGATGATGATGATGATGATGATGATGATG 537
QY 979 TGGGATGACACGAACTTGCCTCCCATGGAACGTGACATGTACTGGAACGAGGACGCC 1038
DB 538 CGAATGTACCAAGAACTTCCCGCGTGAAGAGACAGTACTGGAGATGAGAACAGCC 597
QY 1039 TCAGCCGCCCTTACAC--AGTGTCTTGGCCAGGTCTGCGCGGCAATCTTCAATGGGCTC 1095
DB 598 GCAGCCACCTTACACTGTGCGCCAGCGCGGCGAGTCTGCGGCCCTCATTCCTGGGCTC 657
QY 1096 CACCTTCAACATCAGCTTAAGAAAGA 1122
DB 658 CACCTTCAACCTGCGCATGAGCAGCA 684
```

RESULT 11

US-10-198-846-11070
Sequence 11070, Application US/10198846
Publication No. US2003009974A1
GENERAL INFORMATION:
APPLICANT: Lillie, James
APPLICANT: Xu, Yongyao
APPLICANT: Wang, Youzhen
APPLICANT: Steinhilber, Kathleen
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
FILE REFERENCE: MRI-049
CURRENT APPLICATION NUMBER: US/10/198,846
PRIOR FILING DATE: 2002-07-18
PRIOR APPLICATION NUMBER: 60/306,220
PRIOR FILING DATE: 2001-07-18
NUMBER OF SEQ ID NOS: 14084
SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 11070
LENGTH: 1292
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1287, 1289, 1290, 1291, 1292
OTHER INFORMATION: n = A,T,C or G
US-10-198-846-11070

Query Match 9.5%; Score 181.2; DB 14; Length 1292;
Best Local Similarity 61.4%; Pred. No. 5.9e-40;
Matches 291; Conservative 0; Mismatches 183; Indels 0; Gaps 0;

```
QY 255 TGGGTTCTATGTTACATGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 314
DB 318 TAGGTTTATGTTATCTGTTAGTGAACGATGATGATGATGATGATGATGATGATGATG 377
QY 315 GGGCCGACCGCTCATGATCCAGGTGTCTGATGCTTCTGATGAGGCGAAGATGAGAGGCC 374
DB 378 GGGCAGACAGGCTAATGTTCTCTATCTAGCAATGTTCAAGGAGCGACAGCAGGCGC 437
QY 375 GTTGTGCGGCGGACGCTCATCCGCTACGCAATCTTGGGCGCAAGTGTCTATCTGCGCA 434
DB 438 GCTGCTTGAAGGACCGCTGATGCTACGCTCAATCTTCACTTCTCTCTCTCTCTCTCT 497
QY 435 GCATCAGACCTGCTGCTACAAAGGCTTCCCACTTTCACACCTGCTGCTAGAGGTT 494
DB 498 CGTGAAGACTGCTGTGTACAAAGATTTCCCAATGACACAGCTGTTAAGCAGGTT 557
QY 495 TTATGACCCATGGGACATTAAGCATGTCAGAAAGTTGGGCTTCCACACACACATTTCT 554
DB 558 TTATGACACAGATGAAGAAATTTATTAACACCTCAAGTCTCTCATCTGAATATTT 617
QY 555 GGGTGCCTGGGTGTGTTGGTCCCACTTGTGAATGAAGGCTTATCTTGGAGTGAATCC 614
DB 618 GGGTTCATTCATCTGTTGTTGAATCTTGCACCTAAAGCCGGAATGAAGTGAATCA 677
QY 615 GGGACACCGCTGCTGCTCAGAGCTGATGATGATGATGATGATGATGATGATGATGATG 674
DB 678 GAGACAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 737
QY 675 GACAGCTGTATGCTTACGACTGATGAATATCCATGTTGTGTACACAGGTGG 728
DB 738 GCTCTTATTCGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 791
```

RESULT 12

US-10-027-632-134530/c
Sequence 134530, Application US/10027632
Publication No. US2002019837A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
TITLE OF INVENTION: Polymorphisms in the Human Genome
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
PRIOR FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720


```
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 134530
/ LENGTH: 578
/ TYPE: DNA
/ ORGANISM: Human
US-10-027-632-134530
```

```
Query Match      6.5%; Score 125.2; DB 13; Length 578;
Best Local Similarity 85.8%; Pred. No. 2.4e-24;
Matches 139; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
```

```
QY 488 GCAGGTTTATGACCCATGCGGGAACATTAAGCAGTTGCAAGATTGGGCTTACACACAC 547
    |||
DB 183 GCAGGCTTTATGACTCCGCGAGAACACAGCAGTTGGAGAACTGAGCCTACACACAC 124

QY 548 ACATCTGGGGTCCCTGGGGTGTGGTTTCCAACTGTCAATGAAGAGCCTATCTGGAGGT 607
    |||
DB 123 ATGTTCTGGGGTCCCTGGGGTGTGGTTTCCAACTGTCAATGAAGAGCCTATCTGGAGGT 64

QY 608 CGAATCCGGGACACCGTCTGCTCCAGAGCCTGATGAATGAG 649
    |||
DB 63 CGAATCCGGGACACCGTCTGCTCCAGAGCCTGATGAATGAG 22
```

```
RESULT 13
US-10-027-632-134530/c
/ Sequence 134530, Application US/10027632
/ Publication No. US20030204075A9
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
/ FILE REFERENCE: 108827.129
/ CURRENT APPLICATION NUMBER: US/10/027,632
/ CURRENT FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/198,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ PRIOR FILING DATE: 1999-08-09
/ NUMBER OF SEQ ID NOS: 325720
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 134530
/ LENGTH: 578
/ TYPE: DNA
/ ORGANISM: Human
US-10-027-632-134530
```

```
Query Match      6.5%; Score 125.2; DB 17; Length 578;
Best Local Similarity 85.8%; Pred. No. 2.4e-24;
Matches 139; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
```

```
QY 488 GCAGGTTTATGACCCATGCGGGAACATTAAGCAGTTGCAAGATTGGGCTTACACACAC 547
    |||
DB 183 GCAGGCTTTATGACTCCGCGAGAACACAGCAGTTGGAGAACTGAGCCTACACACAC 124

QY 548 ACATCTGGGGTCCCTGGGGTGTGGTTTCCAACTGTCAATGAAGAGCCTATCTGGAGGT 607
    |||
DB 123 ATGTTCTGGGGTCCCTGGGGTGTGGTTTCCAACTGTCAATGAAGAGCCTATCTGGAGGT 64

QY 608 CGAATCCGGGACACCGTCTGCTCCAGAGCCTGATGAATGAG 649
    |||
DB 63 CGAATCCGGGACACCGTCTGCTCCAGAGCCTGATGAATGAG 22
```

```
RESULT 14
US-10-029-386-7583/c
/ Sequence 7583, Application US/10029386
/ Publication No. US20030194704A1
/ GENERAL INFORMATION:
/ APPLICANT: Penn, Sharon G.
/ APPLICANT: Rank, David R.
/ APPLICANT: Hanzel, David K.
/ TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GI
```

```
FILE REFERENCE: ABOMICA-X-2
CURRENT APPLICATION NUMBER: US/10/029,386
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 34288
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 7583
LENGTH: 539
TYPE: DNA
ORGANISM: Homo sapiens
```

```
FEATURE:
OTHER INFORMATION: MAP TO CHR19.3
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.2
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
OTHER INFORMATION: EST HUMAN HIT: AA573517.1, EVALUATE 9.00e-56
OTHER INFORMATION: NT HIT: G115304094, EVALUATE 1.00e-130
OTHER INFORMATION: SWISSPROT HIT: O76090, EVALUATE 2.00e-25
US-10-029-386-7583
```

```
Query Match      6.4%; Score 122.2; DB 16; Length 539;
Best Local Similarity 68.4%; Pred. No. 1.6e-23;
Matches 169; Conservative 0; Mismatches 78; Indels 0; Gaps 0;
```

```
QY 247 CTTCGTTCTGGGTTTCTATGTTTACATTGGTGTGAGCCGCTGTGTGAGCCAGTACAGAA 306
    |||
DB 372 CTCGCCGACAGGCTTTTATGACGCTGTGTGAGAACCGCTGTGTGAGCCAGTACAGTATG 313

QY 307 CTTCGCGTGGCCGACCGCCTCATGATCCAGGTGTCTAGCTTGTGTGAGGAGCAAGATGA 366
    |||
DB 312 CATGCCCTGCCGACCGCTCATGTGCGTGTGTGAGGAGCAAGCTGTGTGAGGAGCAAGATGA 253

QY 367 GGAAGCGGTTTCTGTGGGCGACCGCTCATCGCTACGCGCATCTGAGGCGCAAGTGTCTAT 426
    |||
DB 252 CCGGCGCGCCTCTACCGGCGGACACTCATGCGCTACGCGCTCTCGGCGTGTCTAT 193

QY 427 CTTCGCGAGCATAGACACTCGGTCTACCAAGCGCTTCCACTCTTCACTCACTGAGTGTCT 486
    |||
DB 192 CTTCGCGTCCGTGACGACCGCGGTGTGTCAAGCGCTTCCCACTCATGACCACTGAGTGTGGA 133

QY 487 AGCAGGT 493
    |||
DB 132 GGCTGGT 126
```

```
RESULT 15
US-10-029-386-21283/c
/ Sequence 21283, Application US/10029386
/ Publication No. US20030194704A1
/ GENERAL INFORMATION:
/ APPLICANT: Penn, Sharon G.
/ APPLICANT: Rank, David R.
/ APPLICANT: Hanzel, David K.
/ TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GI
FILE REFERENCE: ABOMICA-X-2
CURRENT APPLICATION NUMBER: US/10/029,386
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 34288
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 21283
LENGTH: 235
```

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR19.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
; OTHER INFORMATION: EST HUMAN HIT: AA573517.1, EVALUATE 4.00e-56
; OTHER INFORMATION: NT HIT: g18923136, EVALUATE 1.00e-129
; OTHER INFORMATION: SWISSPROT HIT: O76090, EVALUATE 1.00e-24
US-10-029-386-21283

Query Match 6.3%; Score 119.8; DB 16; Length 235;
Best Local Similarity 69.4%; Pred. No. 4.8e-23;
Matches 163; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 258 GTTTCATATGTTACATGTTGTTGTTGAGCCGCTGTTGAGCCAGTACGAGAACTTGCCGTGGC 317
Db 235 GCTTTTATGTAGCGCTGTTGTTGTAACCGCTGTTGAGCCAGTACCTATGTCATGCCGTGC 176
QY 318 CCGACCCGCTCATGATCCAGGTGTTCTAGCTTCGTGAGAGGCAAGATGAGAGGCCGTT 377
Db 175 CCGACGCGCTCATGTCGTGTTGCGGACCGTGCAGACGCGAGCCGCGCGCC 116
QY 378 TGCTGCGGCGCACCGCTATCCGCTACGCGCATCTGGGCCAAGTGCTCACTCTGCGCAGCA 437
Db 115 TCTACCGGCGCACCTCATGCGCTACGAGGGCTCTCGGCCGTGCTCATCTTGCGCTCCG 56
QY 438 TCAGCACCCTCGGCTTACAAAGCGCTTTCCCACTTTTCAACCACTGATGCTAGCAGG 492
Db 55 TCAGCACCAGCGGTTGTTCAAGCGCTTCCCAACATAGACCAAGTGTGAGGCTGG 1

Search completed: March 26, 2005, 22:07:50
Job time : 1075.29 secs

This Page Blank (usp10)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using bw model

Run on: March 25, 2005, 22:47:44 ; Search time 305.441 Seconds
(without alignments)
10264.201 Million cell updates/sec

Title: US-09-622-964a-28
Perfect score: 1916
Sequence: 1 gtcccaagcatgactcatca.....aaaaaaaaaaaaaaaa 1916

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries.

Database : Issued_Patents_NA: *
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	871.2	45.5	2171	4 US-09-949-016-4901	Sequence 4901, Ap
2	188.4	9.8	18075	4 US-09-949-016-16643	Sequence 16643, A
3	125.2	6.5	601	4 US-09-949-016-173051	Sequence 173051, A
4	125.2	6.5	601	4 US-09-949-016-173052	Sequence 173052, A
5	100	5.2	419	4 US-09-270-767-735	Sequence 735, App1
6	100	5.2	419	4 US-09-270-767-16017	Sequence 16017, A
7	72.2	3.6	601	4 US-09-949-016-173045	Sequence 173045, A
8	68.6	3.6	7609	4 US-09-949-016-16644	Sequence 16644, A
9	61.4	3.2	601	4 US-09-949-016-173053	Sequence 173053, A
10	60.4	3.2	601	4 US-09-949-016-173054	Sequence 173054, A
11	52.2	2.7	1361	4 US-09-489-947-64	Sequence 64, App1
12	51.8	2.7	1931	3 US-09-019-942-2	Sequence 2, App1
13	51.8	2.7	1931	3 US-09-099-041A-1	Sequence 1, App1
14	51.8	2.7	1931	3 US-09-245-281-1	Sequence 2, App1
15	51.8	2.7	1931	3 US-09-470-271-2	Sequence 1, App1
16	51.8	2.7	1931	3 US-09-207-359B-1	Sequence 1, App1
17	51.8	2.7	1931	4 US-09-340-620A-1	Sequence 1, App1
18	51.8	2.7	1931	4 US-09-865-364-1	Sequence 1, App1
19	51.8	2.7	1931	4 US-09-748-537-2	Sequence 2, App1
20	51.4	2.7	2619	3 US-08-983-502-17	Sequence 1, App1
21	51.4	2.7	2619	4 US-09-516-747-17	Sequence 1, App1
22	51.4	2.7	2619	5 PCT-US96-10521-17	Sequence 17, App1
23	51.4	2.7	2887	4 US-08-983-502-14	Sequence 14, App1
24	51.4	2.7	2887	3 US-08-516-747-14	Sequence 14, App1
25	51.4	2.7	2887	5 PCT-US96-10521-14	Sequence 14, App1
26	51	2.7	2060	4 US-09-345-473B-5	Sequence 5, App1
27	50.8	2.7	1446	4 US-09-593-359-1	Sequence 1, App1

28	50.6	2.6	1582	3 US-08-545-196B-10	Sequence 10, App1
29	50.6	2.6	1582	4 US-08-545-196B-12	Sequence 12, App1
30	50	2.6	601	3 US-09-949-016-173056	Sequence 173056, A
31	50	2.6	1462	4 US-09-949-016-4902	Sequence 4902, Ap
32	50	2.6	2171	4 US-09-949-016-4901	Sequence 4901, Ap
33	50	2.6	7609	4 US-09-949-016-16644	Sequence 16644, A
34	50	2.6	18075	4 US-09-949-016-16643	Sequence 16643, A
35	49.4	2.6	1740	4 US-09-709-103-45	Sequence 45, App1
36	49.4	2.6	1801	4 US-09-439-410A-45	Sequence 45, App1
37	49.4	2.6	1801	4 US-09-709-103-3	Sequence 3, App1
38	49.4	2.6	1801	4 US-09-439-410A-3	Sequence 3, App1
39	48.6	2.5	1937	4 US-09-647-143-1	Sequence 1, App1
40	48.4	2.5	1584	4 US-09-807-258-11	Sequence 11, App1
41	47.8	2.5	2053	4 US-09-023-655-854	Sequence 854, App
42	47.6	2.5	1046	1 US-08-361-467B-4	Sequence 4, App1
43	47.6	2.5	1046	1 US-08-484-332C-4	Sequence 4, App1
44	47.4	2.5	458	1 US-08-524-757-1	Sequence 1, App1
45	47.4	2.5	1493	1 US-08-340-820-24	Sequence 24, App1

ALIGNMENTS

```
RESULT 1
US-09-949-016-4901
Sequence 4901, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: CU001307
CURRENT APPLICATION NUMBER: US/09/949, 016
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 4901
LENGTH: 2171
TYPE: DNA
ORGANISM: Human
US-09-949-016-4901
Query Match 45.5%; Score 871.2; DB 4; Length 2171;
Best Local Similarity 77.1%; Pred. No. 1.2e+229;
Matches 1116; Conservative 0; Mismatches 253; Indels 39; Gaps 3;
2 TGCCAGCCATGACTATACCTACACAAAGAGCCCAATGCGCGCTGCTGCTTC 61
80 TGCTGCGCATGACACATCACTACACAAAGAGCGCTAATGCGCGCTTATGCTCTTC 139
62 TCGTCCCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 121
140 TCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 199
122 GCTTCATATATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 181
200 ATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 259
182 CAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 241
260 CAACAGCTGATGCTTGAAGAACTGACTCTGATATGCGACAGCTACATCCAGCTATCCC 319
242 ATATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 301
320 ATTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 379
302 GAGAACTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 361
```

Db	380	GAGAACCTGCGGTGGCCCAACCGCCTTCATGAGCTTGATGGGCTTTCGTGAAAGCAAG	439
OY	362	GATGAGAAAGCCGTTTGCTGCGGCGCAAGCTCATCGCTAACCGCAATCTGGGCAAGTG	421
Db	440	GACGACCAAGGCGCGGTGCTGGCGGCAAGCTCATCGGCTAACGCAACCTGGCAACGTG	499
OY	422	CTCATCTCGGCGAGCATCAGCACTTGGGTCTAAGAGCTTTTCCACTTTCACCACTGG	481
Db	500	CTCATCTCGGCGAGCGCTCAGCAACCGAGCTTCAAGAGCTTTCCCGAGCGCCAGCACTGG	559
OY	482	GTGCTTGCAGAGTTTATGAAACCCATGGGAGAAATPAAACAAGTTCAGAAAGTTGGGCTAACCA	541
Db	560	GTGCCAAGCAGGCTTTATGACTCCGCGAAGAACAAACAGATTGGAGAAACTGAGCTTACCA	619
OY	542	CACAAACAAATTCGGGTGCCCTGGAGTGGTTGCCAACTTGTCAATGAAAGCTTACTT	601
Db	620	CACAAACATGTTGGGTGCCCTGGAGTGGTTGCCAACTTGTCAATGAAAGGTGGCTT	679
OY	602	GGAGGTCTGAATCCGGGACACCGTCTGCTTCAGAGCTTGATGATAGGTGTGACTTGG	661
Db	680	GGAGGTCTGAATCCGGGACCCCTATCCCTGCTTCAGAGCTTGTGAACAGATGAAACACTTGG	739
OY	662	CGTACTCAGTGTGAGCAGCTGTATGCTTACAGCTGATTAAGTATCCCATTTGGTGTACACA	721
Db	740	CGTACTCAGTGTGAGCAGCTGTATGCTTACAGCTGATTAAGTATCCCATTTGGTGTATACA	799
OY	722	CAGGTGTGACAGTGGCAGTATACAGCTTTTTCCTTGACAGTGTGATTCGGGAGCGACTTT	781
Db	800	CAGGTGTGACAGTGGCGGTGTACAGCTTCTTCCGATCTGTCTAGTTGGGGGGGAGTTT	859
OY	782	CTGAAACCCAAACAAGACCTAACCGGCACTGAGATGATCTGGTTGTGCTGTCTTACACA	841
Db	860	CTGAAACCCCAAGGCGCTTACCGCTGGCCATGAGCTGGAACCTGTGTGTGCGCTTTCACG	919
OY	842	ATCTGCAATTTCTTATTTACATGGGCTGAGGAGGTGGAGAAACAGCTCATCAACCC	901
Db	920	TTCCTGCAATTTCTTCTTATGTTGGCTGAGAGGTGGAGAGCAGCTCATCAACCC	979
OY	902	TTGCGGAGAGCAGATGATATTTTGGACTTAACTGATCATTTGACAGAAACCTGCAAGTG	961
Db	980	TTTGGAGGAGATGATGATATTTTGGACCAACCTGATTTGTGACAGGAAATTTGGCAGGTG	1033
OY	962	TCCCTGTTGTCCGTGATGAGATGCAACCAAGCTTGCTCCCATGGAACTGACATGTAC	1022
Db	1040	TCCCTGTTGTCCGTGATGAGATGCAACCAAGCCTTGCTCCGATGGAGCGGACATGTAC	1099
OY	1022	TGGAACGAGGCGGCTCAGCGGCGCTTACAGAGCTTCTGCGCAGGCTTGGCGCGCAT	1083
Db	1100	TGGAATTAAGCCCGAGCACAGCCCTCTACACAGCTGTCTTCCGCAAGTTCCGTGAGCC	1155
OY	1082	TCCCTTCAATGGGCTCCACCTTCAACATCAGCTTAAAGAAAGAACTTATAGAGCTTGTGTCA	1144
Db	1160	TCCCTTATGGGCTCCACCTTCAACATCAGCTTAAAGAAAGAAATGATTCAGAGCC	1215
OY	1142	AAAGGAGAGCTGACACGGATPAGAAAGAGAGTGCCTATAGCAGACCAATAGCTGTCTT	1201
Db	1220	AATCAGGAGAGCAGAGAGGATCTCACGCTG-----GCATCATTTGGCGGCTTC	1267
OY	1202	TTAGAGCTGCAACCCAAACATACATCTTCCCTTGAAGACTTTTAAACCAAACTAATTG	1266
Db	1268	CTAGGCTGCAAGTCCCATGATCACATCTTCCCAAGGCAAACTCAAGACCAAACTACTG	1327
OY	1262	TGTTCTAAGA-----ACCCCTCTCGAAGGCGAGTGAAGATGCCAACAGAAAAC	1313
Db	1328	TGGCCCAAGAGGAAATCCCTTCTCACAGAGGCTTGCCCAAAAACCAACAGGACCCAAA	1387
OY	1316	CAGAAAG-----ATGTCTGAAATTTTAAAGGCTTGGACTTCTTG	1356
Db	1388	CAGAAACCTTAAAGGCGCAGAGAAACAACAGGCTGGAAGCTTAAAGCTGTGGAAGCCTTC	1447
OY	1355	AAATGTGTTCCAGGTTTAAAGAGAGGCTTCCATTGTGGCCACAGGACCCACGACG	1414

Db 1448 AAGTCTGCCCACACTGTATCAGAGGCGAGGCTACTACAGTGCCTCCACAGACGCCCCCTCAGC 1507

Qy 1415 CACCCCTAC 1422
| | | |
1508 CCCACTCC 1515

Db

```

RESULT 2
US-09-949-016-16643
; Sequence 16643, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16643
; LENGTH: 18075
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16643

```

```

Query Match          9.8%; Score 188.4; DB 4; Length 18075;
Best Local Similarity 81.1%; Pred. No. 8.6e-41;
Matches 219; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

Dh
233 CTATCCCTTATATCTTGGTTCCTGGGTTTCTATGTTATATTGATGGAGCCGCTGTGG
7315 CTCGCCCCCGGCCCTCTCTGCCCCAGAGCTTACGTGAAGCGTGTGTGAACCGCGTGTGG
7374

Qy
293 AGCCAGTACGAGAACTTGGCCGTGGCCGACCGCTCATGATCCAGGTGTCTAGCTTCTG
7375 AACACGATACGAAACCTGCGCTGACCCGACCGCTCATGAGACCTGGTGTGGGCTTCTC
7434

Qy
353 GAGGGCAAGATGATGAGAAAGCCGTTTGGCTGCGGCGACGCTCATCGGCAAGCCATCTCG
7435 GAAGGCAAGACGACGACAGAGCCGCGCTGTGCGGCGACGCTCATCGGCAAGCCATCTG
7494

Dh
7495 GGCAACGCTCATCTCTGCGGACGCTGACGACCCGACGCTCAAGCGCTTCCCCAGCGCC
7554

Qy
473 CACCACTGTGTCTAGCAGGCTTTATGACC 502
7555 CAGCACTGTGTGCAAGCAGGTGGGAGACC 7584

RESULT 3
US-09-949-016-173051
; Sequence 173051, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTNER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH OR OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498

```

;; PRIOR FILING DATE: 2000-09-08
;; NUMBER OF SEQ ID NOS: 207012
;; SOFTWARE: FASTSEQ for Windows Version 4.0
;; SEQ ID NO: 173051
;; LENGTH: 601
;; TYPE: DNA
;; ORGANISM: Human
US-09-949-016-173051

Query Match 6.5%; Score 125.2; DB 4; Length 601;
Best Local Similarity 85.8%; Pred. No. 3.3e-24;
Matches 139; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy 488 GCAGGTTTATGACCCATGAGGAGAAACATAGAGTTGCGAAGTTGGGCTTACCAACAC 547
Db 321 GCAGGCTTTATGACTCCCGGACAGAACAGAGATTGGAGAACTGAGCTTACCAACAC 380
Qy 548 ACATTCTGGGTGCTTGGGTGTGTTTCCAACTTGCAATGAAAGCCCTATCTTGAGGT 607
Db 381 ATGTTCTGGGTGCTTGGGTGTGTTTCCAACTTGCAATGAAAGCGTGGCTTGAGGT 440
Qy 608 CGAATCCGGGACACCGTCTCTGCTCCAGAGCCTGATGAATGAG 649
Db 441 CGAATCCGGGACCCATCTCTGCTCCAGAGCCTGCTGAACGTG 482

RESULT 4

US-09-949-016-173052
; Sequence 173052, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OR INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO: 173052
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-173052

Query Match 6.5%; Score 125.2; DB 4; Length 601;
Best Local Similarity 85.8%; Pred. No. 3.3e-24;
Matches 139; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy 488 GCAGGTTTATGACCCATGAGGAGAAACATAGAGTTGCGAAGTTGGGCTTACCAACAC 547
Db 30 GCAGGCTTTATGACTCCCGGACAGAACAGAGATTGGAGAACTGAGCTTACCAACAC 89
Qy 548 ACATTCTGGGTGCTTGGGTGTGTTTCCAACTTGCAATGAAAGCCCTATCTTGAGGT 607
Db 90 ATGTTCTGGGTGCTTGGGTGTGTTTCCAACTTGCAATGAAAGCGTGGCTTGAGGT 149
Qy 608 CGAATCCGGGACACCGTCTCTGCTCCAGAGCCTGATGAATGAG 649
Db 150 CGAATCCGGGACCCATCTCTGCTCCAGAGCCTGCTGAACGTG 191

RESULT 5

US-09-270-767-735
; Sequence 735, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:

;; APPLICANT: Homburger et al.
;; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
;; FILE REFERENCE: File Reference: 7326-094
;; CURRENT APPLICATION NUMBER: US/09/270,767
;; CURRENT FILING DATE: 1999-03-17
;; NUMBER OF SEQ ID NOS: 62517
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO: 735
;; LENGTH: 419
;; TYPE: DNA
;; ORGANISM: Drosophila melanogaster
US-09-270-767-735

Query Match 5.2%; Score 100; DB 4; Length 419;
Best Local Similarity 57.9%; Pred. No. 2.4e-17;
Matches 197; Conservative 0; Mismatches 140; Indels 3; Gaps 1;

Qy 851 TTCTTAATTCATGATGGCTGCTGAGAGTGGCAGAAACGCTCATCAACCCCTTGGGGAG 910
Db 2 TTCTTCTTCAATGATGGGTGGCTCAAGGTGGCCGAGTGTGATTAATTCATTGGCGAA 61
Qy 911 GACGATGATGTTTGGAGACTCACTGATCTTTGACAGAAACCTGCAGGTGCTCCTGTG 970
Db 62 GACGATGATGATTTTGGAGTCACTGATGTGATGCAATCTTCAGGTGCTCTATCTG 121
Qy 971 TCCGATGGAGATGACACAGAACTGGCTCCCATGAACTGACATGATCGAAGCGAG 1030
Db 122 ATGCTGACGAGATGACACATGACATTCGAGCTGTTAAAGATCACTATCTGGAGCAG 181
Qy 1031 GCAGCGCC--TCAGCCGCCCTACACAGCTGCTTTCGCAAGTCTGCGGCAATTCCTTC 1087
Db 182 GTGTTCCCAACGAGCTGCCCTTACACATAGCTGCCGAACGATTCGGGAGAAATCATCA 241
Qy 1088 ATGGCTTCACCTTCAACATGAGCTTAAGAAAGAAAGCTTAGAGCTTTGTCAAAAAG 1147
Db 242 GAGCGGTCTCACTGCAAGTCAAGTGGCCCAAGATGGGCGCATGCAATGATGCG 301
Qy 1148 GAGCGTGAACAGATTAAGAAAGAGTGGCTATGACAGCA 1187
Db 302 TCCGTTGCAATGATGAATGCGCGATGATCCAGTGGCA 341

RESULT 6

US-09-270-767-16017
; Sequence 16017, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 16017
; LENGTH: 419
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-16017

Query Match 5.2%; Score 100; DB 4; Length 419;
Best Local Similarity 57.9%; Pred. No. 2.4e-17;
Matches 197; Conservative 0; Mismatches 140; Indels 3; Gaps 1;

Qy 851 TTCTTAATTCATGATGGCTGCTGAGAGTGGCAGAAACGCTCATCAACCCCTTGGGGAG 910
Db 2 TTCTTCTTCAATGATGGGTGGCTCAAGGTGGCCGAGTGTGATTAATTCATTGGCGAA 61
Qy 911 GACGATGATGTTTGGAGACTCACTGATCTTTGACAGAAACCTGCAGGTGCTCCTGTG 970
Db 62 GACGATGATGATTTTGGAGTCACTGATGTGATGCAATCTTCAGGTGCTCTATCTG 121
Qy 971 TCCGATGGAGATGACACAGAACTGGCTCCCATGAACTGAACTGATCTGGAACGAG 1030

RESULT 10
US-09-949-016-173054
; Sequence 173054, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C0001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 173054
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-173054

Query Match
Best Local Similarity 3.2%; Score 60.4; DB 4; Length 601;
Matches 64; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 663 GTACTCAGTGTGACAGCTGTATGCTTACGACGATAGTATCCATTGCTGTACACAC 722
DB 1 GTACTCAGTGTGACACCTGTATGCTTACGACGATAGTATCCATTGCTGTATACAC 60

QY 723 AGCTGCTGAC 732
DB 61 AGCTGAGGAC 70

RESULT 11
US-09-489-847-64
; Sequence 64, Application US/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
; APPLICANT: ROSEN, et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031P1
; CURRENT APPLICATION NUMBER: US/09/489,847
; CURRENT FILING DATE: 2000-01-24
; EARLIER APPLICATION NUMBER: PCT/US99/17130
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/096,319
; EARLIER FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: 60/095,454
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/095,455
; EARLIER FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 64
; LENGTH: 1361
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-489-847-64

Query Match
Best Local Similarity 2.7%; Score 52.2; DB 4; Length 1361;
Matches 69; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 1820 CTGATTTTACGACTTCCCAACTAAGAGTTTAAATAAACGGAATATTTCTTTAGCT 1879

DB 1216 CAGTTGGAATATTAATTTAGAACTCATGAGATTAACCCATTATTTATTTATTA 1275

QY 1880 GAAAAAATT 1916

DB 1276 AA 1312

RESULT 12
US-09-019-942-2
; Sequence 2, Application US/09019942
; Patent No. 603855
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: GENES ENCODING CASPASE RECRUITMENT
; TITLE OF INVENTION: DOMAIN POLYPEPTIDES
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FASTSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/019,942
; FILING DATE: 06-FEB-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Melkielejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 07334/068001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1931 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-019-942-2

Query Match
Best Local Similarity 2.7%; Score 51.8; DB 3; Length 1931;
Matches 71; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 1814 AACACCTGATTTTACGACTTCCCAACTAAGAGTTTAAATAAACGGAATATTTCTT 1873

DB 1825 AAAAGCATGTAAAGTGTGTTTCAAGAAAGAAATGTTTCATTAAGATATTTATTA 1884

QY 1874 TTAGTGAAAAAATTT 1916

DB 1885 AA 1927

RESULT 13
US-09-099-041A-1
; Sequence 1, Application US/09099041A
; Patent No. 6340576
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
; TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF

FILE REFERENCE: 07334-076001
CURRENT APPLICATION NUMBER: US/09/099,041A
CURRENT FILING DATE: 1998-06-17
PRIORITY APPLICATION NUMBER: 09/019,942
PRIOR FILING DATE: 1998-02-06
NUMBER OF SEQ ID NOS: 37
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 1
LENGTH: 1931
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (214)...(1833)
US-09-099-041A-1

Query Match 2.7%; Score 51.8; DB 3; Length 1931;
Best Local Similarity 68.9%; Pred. No. 0.0011;
Matches 71; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 1814 AACACCTGATTTTACGATCTTCCCAACTAAGAGTTTAAATGAATATTTCTT 1873
DB 1825 AAAAGCATGTAGAGCTGTTTTCAGAGAAATGTGTTTCATTAAGATATTTATA 1884
QY 1874 TTAGGTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1916
DB 1885 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1927

RESULT 14
US-09-245-281-1
Sequence 1, Application US/09245281
Patent No. 6369196
GENERAL INFORMATION:
APPLICANT: Bertin, John
TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY
FILE REFERENCE: 07334/118001
CURRENT APPLICATION NUMBER: US/09/245,281
CURRENT FILING DATE: 1999-02-05
EARLIER APPLICATION NUMBER: US 09/207,359
EARLIER FILING DATE: 1998-12-08
EARLIER APPLICATION NUMBER: US 09/099,041
EARLIER FILING DATE: 1998-06-17
EARLIER APPLICATION NUMBER: US 09/019,942
EARLIER FILING DATE: 1998-02-06
NUMBER OF SEQ ID NOS: 44
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 1
LENGTH: 1931
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (214)...(1833)
US-09-245-281-1

Query Match 2.7%; Score 51.8; DB 3; Length 1931;
Best Local Similarity 68.9%; Pred. No. 0.0011;
Matches 71; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 1814 AACACCTGATTTTACGATCTTCCCAACTAAGAGTTTAAATGAATATTTCTT 1873
DB 1825 AAAAGCATGTAGAGCTGTTTTCAGAGAAATGTGTTTCATTAAGATATTTATA 1884
QY 1874 TTAGGTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1916
DB 1885 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1927

RESULT 15
US-09-470-271-2
Sequence 2, Application US/09470271

Patent No. 6410689
GENERAL INFORMATION:
APPLICANT: Bertin, John
TITLE OF INVENTION: GENES ENCODING CASPASE RECRUITMENT
TITLE OF INVENTION: DOMAIN POLYPEPTIDES
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FASTSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/470,271
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/019,942
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Melk1eJohn, Ph.D., Anita L.
REGISTRATION NUMBER: 35,283
REFERENCE/DOCKET NUMBER: 07334/068001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1931 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-470-271-2

Query Match 2.7%; Score 51.8; DB 3; Length 1931;
Best Local Similarity 68.9%; Pred. No. 0.0011;
Matches 71; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 1814 AACACCTGATTTTACGATCTTCCCAACTAAGAGTTTAAATGAATATTTCTT 1873
DB 1825 AAAAGCATGTAGAGCTGTTTTCAGAGAAATGTGTTTCATTAAGATATTTATA 1884
QY 1874 TTAGGTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1916
DB 1885 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1927

Search completed: March 26, 2005, 13:49:44
Job time : 308.441 secs